Quantum Chemistry and Dynamics of Excited States

Quantum Chemistry and Dynamics of Excited States

Methods and Applications

Edited by

Leticia González
Institute of Theoretical Chemistry, Faculty of Chemistry
University of Vienna
Austria

Roland Lindh
Department of Chemistry – BMC
Uppsala University
Sweden

WILEY

This edition first published 2021

Registered Offices
John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA
John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

Editorial Office
The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

For details of our global editorial offices, customer services, and more information about Wiley products visit us at www.wiley.com.

Wiley also publishes its books in a variety of electronic formats and by print-on-demand. Some content that appears in standard print versions of this book may not be available in other formats.

Library of Congress Cataloging-in-Publication Data applied for

ISBN HB: 9781119417750

Cover Design: Wiley
Cover Image: Cover image by Ignacio Fernandez Galvan;
© Thomas Engzell/EyeEm/Getty Images

Set in 9.5/12.5pt STIXTwoText by SPi Global, Chennai, India

10 9 8 7 6 5 4 3 2 1

To our families,
for their unconditional love and endless support.

Contents

List of Contributors

Federica Agostini
Institut de Chimie Physique UMR8000
Université Paris-Saclay, CNRS
Orsay
France

Guillermo Albareda
Max Planck Institute for the Structure and Dynamics of Matter and Center for Free-Electron Laser Science
Hamburg
Germany

Mario Barbatti
Aix Marseille University
CNRS, ICR
Marseille
France

M. Bonfanti
Institute of Physical and Theoretical Chemistry
Goethe University Frankfurt
Frankfurt
Germany

I. Burghardt
Institute of Physical and Theoretical Chemistry
Goethe University Frankfurt
Frankfurt
Germany

Michele Ceotto
Dipartimento di Chimica
Università degli Studi di Milano
Milano
Italy

Riccardo Conte
Dipartimento di Chimica
Università degli Studi di Milano
Milano
Italy

Basile F. E. Curchod
Department of Chemistry
Durham University
Durham
UK

Werner Dobrautz
Department of Electronic Structure Theory
Max Planck Institute for Solid State Research
Stuttgart
Germany

Andreas Dreuw
Interdisziplinäres Zentrum für Wissenschaftliches Rechnen
Ruprecht-Karls Universität Heidelberg
Heidelberg
Germany

Claudia Filippi
MESA+ Institute for Nanotechnology
University of Twente
Enschede
The Netherlands

Jonas Feldt
MESA+ Institute for Nanotechnology
University of Twente
Enschede
The Netherlands

Nicolas Ferré
Aix Marseille University
CNRS, ICR
Marseille
France

Leon Freitag
Institute of Theoretical Chemistry
Faculty of Chemistry
University of Vienna
Vienna
Austria

Ignacio Fdez. Galván
Department of Chemistry – BMC
Uppsala University
Uppsala
Sweden

Sandra Gómez
Institute of Theoretical Chemistry
Faculty of Chemistry
University of Vienna
Vienna
Austria

Leticia González
Institute of Theoretical Chemistry
Faculty of Chemistry
University of Vienna
Vienna
Austria

E. K. U. Gross
Max-Planck-Institut für Mikrostrukturphysik
Halle
Germany

and

Fritz Haber Center for Molecular Dynamics
Institute of Chemistry
The Hebrew University of Jerusalem
Jerusalem
Israel

Kai Guther
Department of Electronic Structure Theory
Max Planck Institute for Solid State Research
Stuttgart
Germany

Erik Donovan Hedegård
Division of Theoretical Chemistry
Lund University
Lund
Sweden

Miquel Huix-Rotllant
Aix Marseille University
CNRS, ICR
Marseille
France

Daniel Keefer
Department Chemie
Ludwig-Maximilians-Universität München
München
Germany

Adam Kirrander
EaStCHEM, School of Chemistry
University of Edinburgh
Edinburgh
UK

Benjamin Lasorne
Institut Charles Gerhardt Montpellier
CNRS – Université de Montpellier
Montpellier
France

Giovanni Li Manni
Department of Electronic Structure Theory
Max Planck Institute for Solid State Research
Stuttgart
Germany

Roland Lindh
Department of Chemistry – BMC
Uppsala University
Uppsala
Sweden

Hans Lischka
Department of Chemistry and Biochemistry
Texas Tech University
Lubbock, Texas
USA

and

School of Pharmaceutical Science and Technology
Tianjin University
Tianjin
P.R. China

Dongxia Ma
Department of Electronic Structure Theory
Max Planck Institute for Solid State Research
Stuttgart
Germany

Sebastian Mai
Institute of Theoretical Chemistry
Faculty of Chemistry
University of Vienna
Vienna
Austria

Philipp Marquetand
Institute of Theoretical Chemistry
Faculty of Chemistry
University of Vienna
Vienna
Austria

Monika Musiał
Institute of Chemistry
University of Silesia in Katowice
Katowice
Poland

Felix Plasser
Department of Chemistry
Loughborough University
Loughborough
UK

Markus Reiher
ETH Zürich
Laboratorium für Physikalische Chemie
Zürich
Switzerland

Sebastian Reiter
Department Chemie
Ludwig-Maximilians-Universität München
München
Germany

Jeremy O. Richardson
Laboratory of Physical Chemistry
ETH Zürich
Zürich
Switzerland

Johan E. Runeson
Laboratory of Physical Chemistry
ETH Zürich
Zürich
Switzerland

Maximilian A. C. Saller
Laboratory of Physical Chemistry
ETH Zürich
Zürich
Switzerland

Ivano Tavernelli
IBM Research GmbH
Zürich Research Laboratory
Rüschlikon
Switzerland

Morgane Vacher
Department of Chemistry – Ångström
The Theoretical Chemistry Programme
Uppsala University
Uppsala
Sweden

Regina de Vivie-Riedle
Department Chemie
Ludwig-Maximilians-Universität München
München
Germany

Graham A. Worth
Department of Chemistry
University College London
London
UK

Preface

Light is not everything, but without light there would not be life as we know it. For good reasons, the sun has been venerated in most ancient civilizations as a god. It has a powerful light and affects our lives on a daily basis. Light drives, controls, or is generated in numerous chemical reactions in nature. It is responsible for processes as essential as vision or photosynthesis, is part of the enchanting phenomena which is bioluminescence but can also be detrimental, such as causing skin damage. Its power has been exploited by mankind since earliest times, not only to heat but to heal, e.g., treating epidermal conditions, to name just one example. Today, sunlight is the hope for providing clean renewable energies. The ways in which the benefits of light can be harnessed go beyond the boundaries of chemistry, into physics, biology and medicine.

Driven by curiosity and interest, many researchers have been fascinated in understanding how light interacts with molecules as only then light can be rationally exploited in many applications. Theory is particular useful for this endeavor, as it allows many details, which are often invisible in experiments, to be disentangled. This particular research field, that could be termed as theoretical photochemistry, took off in the early nineties. Thus, the field is not new but there is still much room for further developments. When a molecule receives a photon of light, its energy changes, or in the language of quantum mechanics, it gets electronically excited. Thus, the computational study of electronic excited states requires the inclusion of quantum effects (at least in part) and this makes it still a challenging problem today except for the smallest molecules. However, recent years have witnessed an explosion of methods able to tackle the study of electronic states and its evolution in time in many different ways. This expansion has been accompanied by thousands of publications dealing with applications involving light. As of 2020, a search in the Web of Science with the words “excited states” and “theory” returned almost 30 000 hits, of which half are just from the last ten years! It is for this reason, that we considered it appropriate to bring this book to light(!), introducing advanced undergraduates, graduate students, and interested researchers to the many flavors in which the field has developed so far.

The book was born with the ambition to collect most of the computational methods that exist today able to solve first the time-independent and then the time-dependent Schrödinger equation for electronic excited states. Accordingly, after an introductory chapter dealing with basic concepts, the book is divided into two parts. Part I contains 9 chapters dealing with electronic structure theory, i.e., solving the time-independent Schrödinger equation and creating building blocks to be used in subsequent dynamics simulations. Part II is divided in 10 chapters devoted to the dynamics of molecules, i.e., solving the time-dependent equation. The solution to the former equation provides energies and other properties, of the electronic excited states in a static manner. It provides multidimensional potential energy surfaces and the corresponding wave functions associated to a particular geometry, which allow molecular spectroscopical properties to be computed. The latter

equation delivers a complementary picture of the system, being in motion, indulging time scales and predicting branching ratios. The richness and broadness of the book invites the reader to reflect on which method could be suitable for a particular problem. We hope that this book fills a gap in the theoretical and computational community dealing with light–matter interactions and becomes a guide in hand, as well as a reference for scientists in the field.

From basic theoretical foundations to the latest theoretical developments, every chapter is self-contained and encompasses the fundamental ideas behind a particular method, its strengths and limitations, as well as selected applications. The chapters are written with the aim to be understandable by master students and newcomers to the field while also informing experts about the state-of-the-art in the field.

Last but not least, we want to express our warm gratitude to all the authors who gracefully accepted the invitation to be part of this adventure, for their enthusiasm, patience and critical suggestions. We are also thankful to our coworkers, for lively discussions and helpful exchanges, in particular, to Philipp Marquetand, Sebastian Mai, Sandra Gómez, and Ignacio Fernández Galván. Their help was priceless.

Leticia González (Vienna)
Roland Lindh (Uppsala)

1

Motivation and Basic Concepts

Sandra Gómez[1], Ignacio Fdez. Galván[2], Roland Lindh[2], and Leticia González[1]

[1] *Institute of Theoretical Chemistry, Faculty of Chemistry, University of Vienna, Vienna, Austria*
[2] *Department of Chemistry – BMC, Uppsala University, SE-751 23 Uppsala, Sweden*

Abstract

This chapter describes what electronic excited states are and why they are important to study and therefore motivates the need for theoretical tools able to characterize them. Further and most importantly, in this introductory chapter, we put together in a comprehensive manner a collection of basic concepts that might be needed, depending on the background of the reader, to understand the remaining chapters of this book.

1.1 Mission and Motivation

When a photon of light strikes a molecule, the latter's electrons are promoted from the electronic ground state to higher electronic levels. Typically, the electronic ground state of a molecule is a singlet state, but depending on the number of electrons and their most favorable way of pairing, it can be a doublet, a triplet, or a state of higher multiplicity. Assuming the electronic ground state is a singlet, upon light absorption the molecule will be excited to another singlet state, as high in energy as the energy contained in the photon allows. Once excited, a number of radiative and non-radiative decay processes are possible. These are collected in the Jabłoński diagram shown in Figure 1.1(a), which assumes an electronic singlet ground state.

Radiative processes include fluorescence or phosphorescence, depending on whether the emission of light involves a transition between two states of the same multiplicity, for example from the lowest singlet S_1 to the S_0, or involves a change of spin, as shown in Figure 1.1, from the triplet T_1 to the S_0. Typically, as in the example depicted, the emitted light has a longer wavelength than the absorbed radiation because luminescence occurs from lower energy levels, and thus absorption and emission spectra are easy to identify from experimental data. In this example, the molecule returns to the original ground state from where it started and thus there was no photochemical reaction, one would say that a photophysical process has taken place.

Non-radiative processes can be much more complicated to observe experimentally, as they typically involve not only the bright or absorbing state defined by the wavelength employed to irradiate, but also dark states, i.e., states that do not have a significant oscillator strength but are populated from the bright states. A transition between electronic states of the same multiplicity is known as

Quantum Chemistry and Dynamics of Excited States: Methods and Applications,
First Edition. Edited by Leticia González and Roland Lindh.

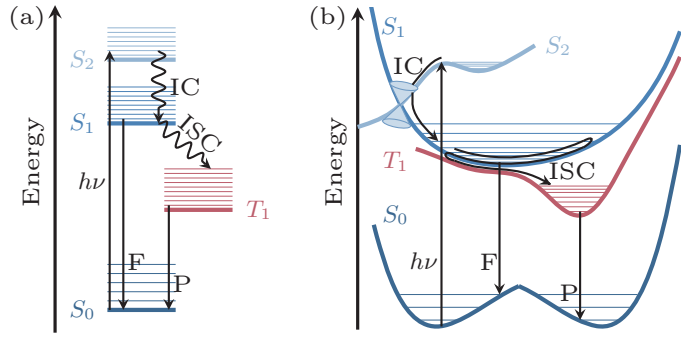


Figure 1.1 (a) Jabłoński diagram with levels. After absorption of a photon with energy $h\nu$, different processes can occur: radiative processes are fluorescence (F) and phosphorescence (P), non-radiative processes are internal conversion (IC) and intersystem crossing (ISC). (b) Jabłoński diagram with potential energy surfaces.

internal conversion, e.g., from S_2 to S_1. When two states of different multiplicities are involved, e.g., from the S_1 to T_1, one speaks of intersystem crossing.

The electronic levels of a molecule are defined through potential energy surfaces (PES) that extend along $3N - 6$ dimensions (with N the number of atoms contained in the molecule). PES are the direct consequence of invoking the Born–Oppenheimer approximation (BOA), see section 1.7. As comfortable as it might seem for a chemist to employ electronic states to envision the course of a chemical reaction from a reactant to a product, sticking to the BOA when talking about electronic excited states implies that the coupling between different PES is neglected. However, these so-called *non-adiabatic couplings* between PES are the "salt and pepper" of photochemistry, as they are essential to understand which states and geometrical conformations are populated after excitation. One key concept in this respect is the non-adiabatic transition around a *conical intersection*, see section 1.9. Named after the ideal topology two PES adopt when they intersect (see Figure 1.1(b)), a conical intersection is the molecular funnel that allows for internal conversion, and it can also be seen as the transition state in photochemistry, which connects a reactant with a product. Likewise, intersystem crossing is mediated by spin–orbit coupling, which is another form of vibronic or non-adiabatic coupling between electronic levels.

Figure 1.1(b) summarizes the radiative and non-radiative processes described before, now in terms of PES. If after the detour via the different PES, the molecule ends up at a different geometrical configuration from which it started after irradiation, one speaks of a photochemical reaction; if instead, it returns back to the electronic ground state of the reactant, the term photophysics is employed.

Be it photophysics or photochemistry, light-induced processes are all around us. As Ciamician already recognized in 1912[1], "reactions caused by light are so many, that it should not be difficult to find some of practical value". Indeed, just to give one representative example, the dream of using solar fuel to produce sustainable energy is keeping many scientists around the world busy. In an effort to mimic natural photosynthesis, one needs among others, to design efficient antenna complexes able to harvest the broad solar spectrum and direct the electrons towards the catalytic centers. This design requires a profound understanding of the underlying processes that take place in the molecules after light excitation. Theoretical modeling can help explain existing experiments and hopefully guide new ones. Which are the electronic states that are populated after

1 Giacomo Ciamician, "The photochemistry of the future", *Science* 36 (1912) 385–394.

excitation? How does the molecule evolve along the complicated PES associated to these electronic states? Often these two simple questions are not easy to answer. They imply a need to get an accurate solution of two key equations, the electronic time-independent Schrödinger equation and the time-dependent Schrödinger equation. Both equations are challenging to solve, except for very small molecules, and so approximations and numerical strategies are required. The solution of the first equation is the goal of electronic structure theory and the solution of the second, the target of chemical dynamics. Both fields have tremendously evolved in the last decades, with the emergence of many different methods that have a common objective.

The mission of this book is to keep up-to-date with the recent development in these two intertwined fields, setting the focus at solving electronic excited states and following their time evolution. Accordingly, Part I collects the most important electronic structure methods that can be used nowadays to calculate electronic excited states as well as associated PES and other electronic properties. Part II, in turn, covers the state of the art for solving molecular motion in the electronic excited states. The variety and extension of the methods collected in this book speaks for itself about how much progress has been achieved in this branch of theoretical chemistry, which undoubtedly has also massively profited in the last years from enormous advances in computational resources. It would not be fair, however, to pretend that theoretical photochemistry has reached its cusp. A deeper reading of the chapters will reveal to the reader not only how far we have come but also how much still remains to be done.

In an effort to make the contents of this book accessible to undergraduates and newcomers to the field, the rest of this chapter contains a number of basic concepts to ease the reading. All the chapters have been written in a fully consistent manner, so as to allow them to be studied independently from the others. The chapters are, nevertheless, organized such that they try to reflect a natural progression. In this respect, the chapters are grouped in two sections consisting of Part I and Part II – electronic structure theory and methods for molecular dynamics, respectively.

In the electronic structure section the selected order of the chapters tries, to some extent, to be in the order of sophistication. However, in some cases chapters are clustered together because of common grounds or methodology. In that sense, Part I starts with the chapters based on density functional theory (DFT) – the chapters on time-dependent DFT (TD-DFT) and multi-configurational DFT (MC-DFT). This is followed by chapters revolving around equation-of-motion coupled cluster theory (EOM-CC) and the algebraic-diagrammatic construction (ADC) scheme for the polarization propagator, which are grouped together due to the technical similarities of the methods. Finally, five chapters are grouped together based on the use of a configurational interaction (CI) type of wave function. Initially, the basics of the so-called complete active space SCF (CASSCF) and related methods – the foundation of multi-configurational quantum chemistry – is introduced. This is followed by two chapters on techniques describing how to solve the associated equations – the chapters on density matrix renormalization group (DMRG) and the quantum Monte-Carlo (QMC) approaches. To conclude Part I, two chapters about the inclusion of electronic dynamical correlation follow – the chapters on the multi-reference configuration interaction (MRCI) method and the multi-configurational reference perturbation theory (MRPT). A pictorial summary of the methods described is provided in Figure 1.2. Starting from Hartree–Fock (HF), different methods cover different degrees of dynamic and static correlation, all the way to the exact full-CI (FCI).

Part II, dealing with the time evolution of nuclear configurations, starts with three chapters that can be considered within the realm of quantum dynamics. The first one introduces the time-dependent Schrödinger equation and how to solve it exactly in a grid – what is known as wave packet dynamics. Due to the cost of obtaining PES, wave packet dynamics is typically done in reduced dimensionality. The multi-configuration time-dependent Hartree (MCTDH) family of

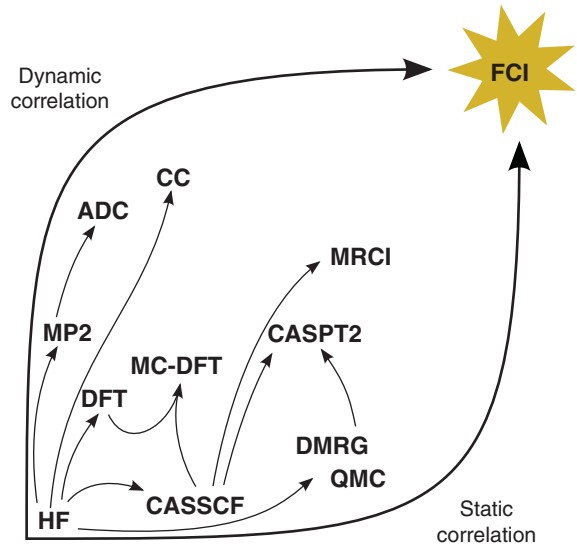


Figure 1.2 Scheme of quantum chemical methods for electronic structure. The lower left corner contains the most basic *ab initio* method, Hartree–Fock (HF), while the exact solution of the time-independent Schrödinger equation, full configuration interaction (FCI), lies, mostly unreachable, on the upper right corner. A panoply of methods described in Part I of this book, identified by their acronyms, try to "correct" HF, adding the missing *electronic correlation* and thus approximating the ideal FCI. The methods are arranged, qualitatively, based on their algorithmic relations and their prioritization of so-called *static* or *dynamic* correlation, which ultimately lead to the same end point.

methods is presented next, as a method that can alleviate in part the cost of grid-based wave packet methodologies. This chapter ends, bridging with the next block of four chapters that are based on quantum-mechanical and quantum-classical methods using on-the-fly computation of PES. These chapters are arranged in sort of going from more to less "quantum" – direct dynamics variational multi-configurational Gaussian (DD-vMCG) method, full and ab initio multiple spawning (FMS and AIMS), Ehrenfest methods, and surface hopping (SH). The next four chapters are based on alternative formulations of quantum dynamics. Exact factorization is based on an alternative way to express the electronic–nuclear wave function, Bohmian dynamics is based on wave theory, while semi-classical and path integral methods are based on Feynman's path integral formulation. Figure 1.3 illustrates pictorically the dynamical methods explained here.

Given the diversity of methods and authors it is unavoidable that every chapter follows its own writing style. For that reason, we considered it useful to collect here some underlying mathematical background, assuming basic knowledge of quantum mechanics, as well as a few photochemical concepts, that naturally arise in many chapters.

1.2 Atomic Units

A comment on atomic units is in order here. Hartree atomic units can elegantly simplify equations by setting to 1 the numerical value of some fundamental constants. Typical examples are the mass of the electron m_e, the electron charge e, the Coulomb or electric force constant $k_e = \frac{1}{4\pi\epsilon_0}$ and the reduced Plank constant $\hbar = \frac{h}{2\pi}$. Other useful constants used as units, derived from those fundamental quantities and used in this book are the bohr, $a_0 \approx 0.529$ Å, and the hartree, $E_h \approx 27.21$ eV.

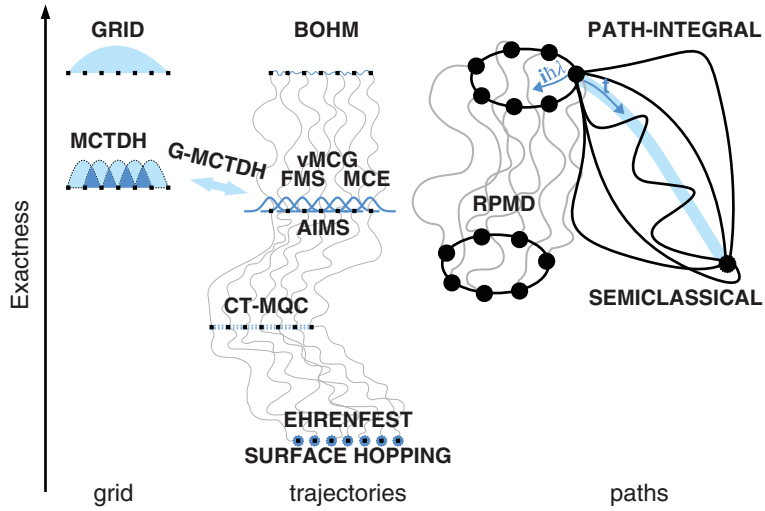


Figure 1.3 Grouping of dynamics methods by families and degree of exactness. Grid methods include the standard method to solve the time-dependent Schrödinger equation (GRID), the multi-configurational time-dependent Hartee method (MCTDH) and partially Gaussian MCTDH (G-MCTDH). From Bohmian trajectories, formally exact to frozen Gaussian methods, such as variational multi-configurational Gaussian (vMCG), full multiple spawning (FMS), multi-configurational Ehrenfest (MCE) and ab-initio multiple spawning (AIMS). Coupled trajectories mixed quantum classical (CT-MQC) is the trajectory method derived from exact factorization. Ehrenfest method and trajectory surface hopping (SH), follow next, since they are based on uncoupled trajectories. Derived from the formally exact path integral method, semi-classical and ring-polymer molecular dynamics (RPMD) arise.

However, using this notation forces readers to keep track of the omitted units, preventing a straightforward dimensionality analysis. For this reason, atomic units have been avoided as much as possible in most chapters, unless otherwise stated.

1.3 The Molecular Hamiltonian

The time evolution of a system is described by the time-dependent Schrödinger equation,

$$\mathrm{i}\hbar\frac{\partial\Psi(\boldsymbol{r},\boldsymbol{R},t)}{\partial t} = \hat{H}\Psi(\boldsymbol{r},\boldsymbol{R},t), \tag{1.1}$$

where $\hat{H}$ is the Hamiltonian of the system and $\Psi(\boldsymbol{r},\boldsymbol{R},t)$ is the wave function describing the molecule, with $\boldsymbol{r}$ and $\boldsymbol{R}$, the electronic and nuclear coordinates, respectively. When applied to the wave function, $\hat{H}$ yields the respective energy. If assumed as time-independent, $\hat{H}$ can be expressed as,

$$\begin{aligned}\hat{H}(\boldsymbol{r},\boldsymbol{R}) = &\underbrace{-\sum_{i=1}^{N}\frac{\hbar^2}{2m_\mathrm{e}}\nabla_i^2}_{\hat{T}^\mathrm{e}} \underbrace{-\sum_{A=1}^{K}\frac{\hbar^2}{2M_A}\nabla_A^2}_{\hat{T}^\mathrm{n}} + \underbrace{\sum_{i=1}^{N}\sum_{j>i}^{N}\frac{e^2}{4\pi\epsilon_0|\boldsymbol{r}_i-\boldsymbol{r}_j|}}_{\hat{V}^\mathrm{ee}} \\ &+\underbrace{\sum_{A=1}^{K}\sum_{B>A}^{K}\frac{e^2Z_AZ_B}{4\pi\epsilon_0|\boldsymbol{R}_A-\boldsymbol{R}_B|}}_{\hat{V}^\mathrm{nn}} \underbrace{-\sum_{i=1}^{N}\sum_{A=1}^{K}\frac{e^2Z_A}{4\pi\epsilon_0|\boldsymbol{r}_i-\boldsymbol{R}_A|}}_{\hat{V}^\mathrm{ne}},\end{aligned} \tag{1.2}$$

where the terms labeled as $\hat{T}$ are kinetic energy terms for K nuclei and N electrons, and the rest, labeled as $\hat{V}$ are the potential energy terms describing interactions between electrons or nuclei themselves, or between nuclei and electrons.

In atomic units, this equation reads simply as

$$\begin{aligned}\hat{H}(\boldsymbol{r},\boldsymbol{R}) = &-\sum_{i=1}^{N}\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{K}\frac{1}{2M_A}\nabla_A^2 + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{|\boldsymbol{r}_i-\boldsymbol{r}_j|}\\ &+\sum_{A=1}^{K}\sum_{B>A}^{K}\frac{Z_A Z_B}{|\boldsymbol{R}_A-\boldsymbol{R}_B|} - \sum_{i=1}^{N}\sum_{A=1}^{K}\frac{Z_A}{|\boldsymbol{r}_i-\boldsymbol{R}_A|}\,.\end{aligned} \tag{1.3}$$

1.4 Dirac or Bra-Ket Notation

The Dirac notation is a very compact way of describing quantum states and their inner products used in quantum mechanics, where the kets ($|\cdot\rangle$) are column vectors and the bras ($\langle\cdot|$) their Hermitian transpose row vectors, such as:

$$\langle\Psi| = |\Psi\rangle^*\,. \tag{1.4}$$

Ket vectors are normally used in this context to specify the state of a system in whatever space basis we are currently working on. Therefore, a wave function Ψ in x coordinates could be expressed as:

$$\Psi(x) \equiv \langle x|\Psi\rangle\,. \tag{1.5}$$

In the same way, applying an operator on this state would lead to:

$$\hat{A}\Psi(x) \equiv \langle x|\hat{A}|\Psi\rangle\,. \tag{1.6}$$

This notation can also be used to express *integrals* over products of wave functions – or rather a wave function and a complex-conjugate wave function. For example, $\langle\Psi|\Psi\rangle$ is nothing but a shorthand notation for

$$\int \Psi^*(\boldsymbol{r})\Psi(\boldsymbol{r})\mathrm{d}\boldsymbol{r} \equiv \langle\Psi|\Psi\rangle\,. \tag{1.7}$$

where $\boldsymbol{r}$ symbolizes all the coordinates on which Ψ depends, and the integration is done over the whole domain. An operator can be included as in

$$\int \Psi^*(\boldsymbol{r})\hat{A}\Psi(\boldsymbol{r})\mathrm{d}\boldsymbol{r} \equiv \langle\Psi|\hat{A}|\Psi\rangle\,, \tag{1.8}$$

In practical quantum chemistry calculations, one commonly uses a one-electron basis set to expand wave functions, i.e., a set of functions that depend on the coordinates of a single electron. In this context, a few particular forms of integrals are especially useful, and their usual notation will be introduced here. First, there is a term that collects the one-electron operators in the Hamiltonian, the kinetic energy and the nuclei–electron attraction (see section 1.3), and is expressed as

$$h_{pq} = \langle\phi_p|\hat{h}|\phi_q\rangle \tag{1.9}$$

$$= -\int \phi_p^*(\boldsymbol{r})\left(\frac{\hbar^2}{2m_\mathrm{e}}\nabla^2 + \frac{e}{4\pi\epsilon_0}\sum_{A=1}^{K}\frac{Z_A}{|\boldsymbol{r}-\boldsymbol{R}_A|}\right)\phi_q(\boldsymbol{r})\mathrm{d}\boldsymbol{r} \tag{1.10}$$

$$= -\frac{\hbar^2}{2m_\mathrm{e}}\int \nabla\phi_p^*(\boldsymbol{r})\nabla\phi_q(\boldsymbol{r})\mathrm{d}\boldsymbol{r} - \frac{e}{4\pi\epsilon_0}\sum_{A=1}^{K}\int\frac{Z_A\phi_p^*(\boldsymbol{r})\phi_q(\boldsymbol{r})}{|\boldsymbol{r}-\boldsymbol{R}_A|}\mathrm{d}\boldsymbol{r} \tag{1.11}$$

where ϕ_p as ϕ_q are two basis functions. The other important term to note corresponds to the two-electron repulsion:

$$(pq|rs) = g_{pqrs} = \langle \phi_p \phi_r | \hat{g} | \phi_q \phi_s \rangle \tag{1.12}$$

$$= \frac{e^2}{4\pi\epsilon_0} \int \frac{\phi_p^*(\boldsymbol{r}_1)\phi_q(\boldsymbol{r}_1)\phi_r^*(\boldsymbol{r}_2)\phi_s(\boldsymbol{r}_2)}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} \mathrm{d}\boldsymbol{r}_1 \mathrm{d}\boldsymbol{r}_2 \,. \tag{1.13}$$

Some alternative notations for these two-electron integrals are:

$$\langle pq|rs\rangle = (pr|qs) \tag{1.14}$$

$$\langle pq||rs\rangle = \langle pq|rs\rangle - \langle pq|sr\rangle. \tag{1.15}$$

1.5 Index Definitions

In several chapters of this book we will use the following convention with respect to the indexation of orbitals;

- $a, b, c, d, \ldots$ to denote empty (virtual) orbitals;
- $i, j, k, l, \ldots$ to denote doubly occupied (inactive) orbitals;
- $t, u, v, x, \ldots$ to denote active orbitals; and
- $p, q, r, s, \ldots$ as general indices to denote orbitals of unspecified type.

1.6 Second Quantization Formalism

We will now give a very brief introduction to the formalism of second quantization and how it is used to express the Hamiltonian. In the original formulation of quantum mechanics for a fixed number of particles, it was natural that the electronic Hamiltonian was expressed in terms of operators which are *sums over particles* – this was later called first quantization,

$$\hat{H}^{\mathrm{el}} = \sum_{i=1}^{N} (\hat{T}_i^{\mathrm{e}}(\boldsymbol{r}_i) + V^{\mathrm{ne}}(\boldsymbol{r}_i)) + \frac{e^2}{4\pi\epsilon_0} \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{1}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} \,, \tag{1.16}$$

where i is the index of an electron, $\hat{T}_i^{\mathrm{e}}$ is the kinetic energy operator of electron i, and

$$V^{\mathrm{ne}}(\boldsymbol{r}_i) = -\frac{e}{4\pi\epsilon_0} \sum_{A=1}^{K} \frac{Z_A}{|\boldsymbol{r}_i - \boldsymbol{R}_A|} \tag{1.17}$$

is the attraction experienced by electron i from all the nuclei in the system. The last term is the electron–electron repulsion term. However, with the onset of quantum field theory the notion of a system having a fixed number of particles was abandoned and a new formalism was introduced – second quantization – in which the *summations run over the orbital space* and the Hamiltonian operator is expressed in terms of operators that "probe" whether a particle (electron) is present in some orbital and if so, include the contribution of that particular orbital. In this formalism the electronic Hamiltonian is expressed as

$$\hat{H}^{\mathrm{el}} = \sum_{pq} \hat{E}_{pq} h_{pq} + \frac{1}{2} \sum_{pqrs} \hat{e}_{pqrs} (pq|rs), \tag{1.18}$$

where the summations are now in terms of the electronic orbitals[2], h_{pq} and $(pq|rs)$ are the one- and two-electron integrals introduced in Section 1.4, where ϕ are normalized molecular orbitals. The remaining operators are the "probing" operators in which the first operator

$$\hat{E}_{pq} = \sum_{\sigma=\{\alpha,\beta\}} \hat{a}^{\dagger}_{p_\sigma} \hat{a}_{q_\sigma} \tag{1.19}$$

is the spin-averaged electron replacement operator ($\hat{a}^{\dagger}_{p_\sigma}$ and $\hat{a}_{q_\sigma}$ are standard creation and annihilation operators for electrons of spin σ in orbitals p and q, respectively), which moves one electron from spatial orbital q to orbital p. The second operator is a two-electron replacement operator, which can be expressed in terms of the one-electron replacement operator as

$$\hat{e}_{pqrs} = \hat{E}_{pq}\hat{E}_{rs} - \delta_{qr}\hat{E}_{ps}\,. \tag{1.20}$$

In this formalism the Hamiltonian is now invariant to the number of particles – in terms of calculations in a finite basis; however, the Hamiltonian is a function of the size of the basis set. This formalism has several advantages and is now the standard in multi-configurational electron structure theory. Let us now briefly explore the probing nature of the operator in Eq. (1.19). In the case $\hat{E}_{pp}$ operates on a closed-shell Slater determinant (SD), constructed from a set of orthonormal orbitals, we will have

$$\hat{E}_{pp}\psi_{\mathrm{SD}} = \begin{cases} 2\psi_{\mathrm{SD}} & \text{if } p \text{ is an occupied orbital} \\ 0\psi_{\mathrm{SD}} & \text{if } p \text{ is an empty orbital,} \end{cases} \tag{1.21}$$

the operator will simply try to remove the electron and then try to put it back – this special case of the electron replacement operator is also called the *occupation number operator*. Since every occupied orbital in a closed-shell SD carries two electrons, we will get an occupation number of two, alternatively, if no electrons are found in orbital p the operator will be a null operator. In general we have that

$$\sum_p \hat{E}_{pp}\Psi = N\Psi, \tag{1.22}$$

where N is the number of particles (electrons) in the system. The following commutation relations apply to the one- and two-electron replacement operators:

$$[\hat{E}_{pq}, \hat{E}_{xy}] = \hat{E}_{py}\delta_{xq} - \hat{E}_{xq}\delta_{py} \tag{1.23}$$

$$[\hat{e}_{pqrs}, \hat{E}_{xy}] = \hat{e}_{pyrs}\delta_{xq} - \hat{e}_{xqrs}\delta_{py} + \hat{e}_{pqry}\delta_{xs} - \hat{e}_{pqxs}\delta_{ry}\,. \tag{1.24}$$

Finally we define the one- and two-particle density matrices as

$$D_{pq} = \langle\Psi|\hat{E}_{pq}|\Psi\rangle\,, \tag{1.25}$$

and

$$\Gamma_{pqrs} = \langle\Psi|\hat{e}_{pqrs}|\Psi\rangle\,. \tag{1.26}$$

It is important to note the four-fold permutational symmetry of the two-electron density matrix:

$$\Gamma_{pqrs} = \Gamma_{rspq} = \Gamma_{qpsr} = \Gamma_{srqp}\,. \tag{1.27}$$

2 We have used a shorthand notation to indicate multiple summations ($\sum_{pq} = \sum_p \sum_q$), and we leave the summation limits implicit

In this perspective we compute the electronic energy of a normalized wave function Ψ with the general expression

$$E = \sum_{pq} D_{pq} h_{pq} + \frac{1}{2} \sum_{pqrs} \Gamma_{pqrs} (pq|rs). \tag{1.28}$$

Finally, this expression can be transformed from a molecular orbital basis to any other basis. In particular, given that the orbitals ϕ are expressed as linear combinations of the one-particle basis functions χ as

$$\phi_p = \sum_{\mu} c_{p\mu} \chi_\mu \,, \tag{1.29}$$

we have that the one- and two-electron density matrices in the one-particle basis function expansion are generated by

$$D_{\mu\nu} = \sum_{pq} c_{p\mu} D_{pq} c_{q\nu} \tag{1.30}$$

and

$$\Gamma_{\mu\nu\gamma\delta} = \sum_{pqrs} c_{p\mu} c_{q\nu} \Gamma_{pqrs} c_{r\gamma} c_{s\delta} \,, \tag{1.31}$$

and we get

$$E = \sum_{\mu\nu} D_{\mu\nu} h_{\mu\nu} + \frac{1}{2} \sum_{\mu\nu\gamma\delta} \Gamma_{\mu\nu\gamma\delta} (\mu\nu|\gamma\delta). \tag{1.32}$$

This expression has a special advantage in so-called direct methods, since the energy can be compiled directly from the integrals as they are generated in the one-particle basis set.

1.7 Born–Oppenheimer Approximation and Potential Energy Surfaces

The BOA is one of the most used approximations in quantum chemistry to solve the time-independent Schrödinger equation. Taking advantage of the fact that protons and neutrons are about 2000 times heavier than electrons, it neglects the kinetic energy of the nuclei. Therefore, at every nuclear position the electrons will feel an average potential depending on where the nuclei are located. This fact allows a separation of the Schrödinger equation into an electronic and a nuclear part. It is then possible to solve the electronic Schrödinger equation for every specific nuclear configuration, in which the Hamiltonian, eigenstates and eigenvalues depend parametrically on the nuclear position,

$$\hat{H}^{\mathrm{el}} \Psi^{\mathrm{el}}(\boldsymbol{r}; \boldsymbol{R}) = E^{\mathrm{el}}(\boldsymbol{R}) \Psi^{\mathrm{el}}(\boldsymbol{r}; \boldsymbol{R}). \tag{1.33}$$

Grouping the remaining terms of the molecular Hamiltonian (Eq. (1.2)), one arrives to the time-independent nuclear Schrödinger equation:

$$\mathrm{i}\hbar \frac{\partial |\Psi^{\mathrm{nuc}}(\boldsymbol{R})\rangle}{\partial t} = \hat{H} |\Psi^{\mathrm{nuc}}(\boldsymbol{R})\rangle = [\hat{T}^{\mathrm{n}}(\boldsymbol{R})) + E^{\mathrm{el}}(\boldsymbol{R})] |\Psi^{\mathrm{nuc}}(\boldsymbol{R})\rangle \,. \tag{1.34}$$

This equation represents nuclei that can move on effective potential surfaces represented by the electronic energies E^{el}. This representation of energies of electrons that depend parametrically on the nuclear coordinates is precisely what is called the potential energy surface (PES). This approximation will be valid whenever nuclei and electrons approximately decouple.

1.8 Adiabatic Versus Diabatic Representations

If the nuclei move extremely slowly, the electronic Hamiltonian will change very slowly with time, since it depends on the value of the nuclear coordinates. This means that if the starting electronic state was an eigenstate of the electronic Hamiltonian at the initial position, it will continue being an eigenstate during its time evolution. In this sense, the BOA is also known as the *adiabatic approximation*, since it assumes that the system behaves all the time "adiabatically", i.e., without changing its electronic wave function. However, there are some cases where the BOA breaks down.

Let us here express the total electronic wave function as a product of the electronic and nuclear ones using all electronic eigenstates (α indices):

$$|\Psi(\boldsymbol{r},\boldsymbol{R})\rangle = \sum_{\alpha} |\Psi_{\alpha}^{\mathrm{el}}(\boldsymbol{r};\boldsymbol{R})\rangle \cdot |\Psi_{\alpha}^{\mathrm{nuc}}(\boldsymbol{R})\rangle\,, \tag{1.35}$$

also known as the Born–Huang expansion. This expression can be inserted into the electronic Schrödinger equation to yield

$$\left(\hat{T}^{\mathrm{n}} + E_{\alpha}^{\mathrm{el}}\right)|\Psi_{\alpha}^{\mathrm{nuc}}(\boldsymbol{R})\rangle + \sum_{\beta} \hat{T}_{\beta\alpha}^{\mathrm{NAC}}|\Psi_{\beta}^{\mathrm{nuc}}(\boldsymbol{R})\rangle = E|\Psi_{\alpha}^{\mathrm{nuc}}(\boldsymbol{R})\rangle\,. \tag{1.36}$$

When comparing Eq. (1.36) and Eq. (1.34), one can see that the difference arises from the $\hat{T}_{\beta\alpha}^{\mathrm{NAC}}$ term, that can be expanded as

$$\hat{T}_{\beta\alpha}^{\mathrm{NAC}} = -\sum_{A=1}^{K} \frac{1}{2M_A}[\langle\Psi_{\alpha}^{\mathrm{el}}|\nabla_A^2|\Psi_{\beta}^{\mathrm{el}}\rangle + \langle\Psi_{\alpha}^{\mathrm{el}}|\nabla_A|\Psi_{\beta}^{\mathrm{el}}\rangle\nabla_A]\,. \tag{1.37}$$

Since neglecting these kinetic coupling terms is the core of the BOA, they can be seen as corrections. While the first term is known as BO diagonal coupling and is normally negligible, the second term – called derivative or non-adiabatic coupling (NAC) – can be rather large in regions where the electronic wave function changes fast with the nuclear coordinates. $\nabla_A|\Psi_{\beta}^{\mathrm{el}}\rangle$ is the gradient of the electronic wave function and gives us the direction where it changes fastest. After projecting it onto $\langle\Psi_{\alpha}^{\mathrm{el}}|$, i.e., calculating its overlap with it, the term can be seen as how much the change of the electronic wave function agrees with another electronic eigenstate. Its extent tells us how likely non-adiabatic events are, its direction the coordinate motions where this change is larger, always for a specific pair of electronic states. Note that this equation is still adiabatic, but corrected with terms corresponding to non-adiabatic events or the breakdown of the Born–Oppenheimer approximation. Thus, a total Hamiltonian might be composed of a Born–Oppenheimer one plus the NAC terms that generate transitions between BO states.

In contrast, we could define a basis where the electronic states do not depend on the nuclear coordinates (this labeling of the electronic wave function is normally called its *character*, so in this case we would say the electronic state keeps its electronic character). This basis can be defined at a particular geometry $\boldsymbol{R}_0$ where the electronic states were eigenstates of the electronic Hamiltonian. Those states are called *diabatic* states and do not diagonalize H^{el} at geometries different from $\boldsymbol{R}_0$,

$$|\Psi_{\alpha}^{\mathrm{el}}(\boldsymbol{R})\rangle = \sum_{i} c_{\alpha i}(\boldsymbol{R})|\Psi_{\alpha}^{\mathrm{el}}\rangle_{\mathrm{dia}}\,, \tag{1.38}$$

where there is no dependence of $|\Psi_{\alpha}^{\mathrm{el}}\rangle_{\mathrm{dia}}$ on the nuclear coordinates $\boldsymbol{R}$. Doing so, we can already neglect the NAC terms, since there is no change of the electronic wave function with the nuclear coordinates and therefore the coupling between electronic states is absorbed in the off-diagonal terms of the Hamiltonian matrix. A scheme depicting adiabatic and diabatic states can be seen in Figure 1.4.

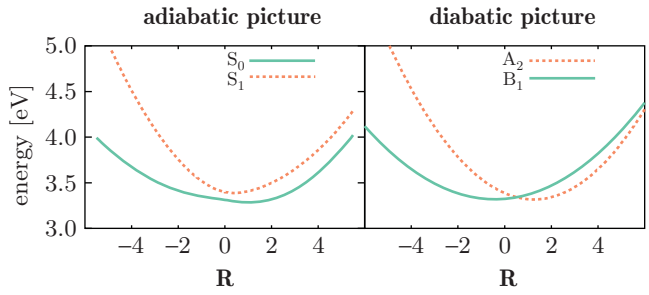


Figure 1.4 Schematic representation of adiabatic (states ordered by energy) and diabatic (states ordered by their symmetry label or character) potential energy curves.

For systems having more than two electronic states, it is usually not possible to find strictly diabatic states, so quasi-diabatic states are defined as a set of electronic states that minimize the NAC terms. In general, the diabatic representation is used whenever one needs to integrate in time domain over all the configurational space (see, for example, chapters 11 and 12) since the diabatic couplings are smoothly varying with the nuclear coordinates, making numerics easier. In contrast, dynamical methods based on local approaches (see e.g., chapters 14, 15 and 16), prefer the adiabatic picture due to the peaked localized NACs that will tell when the BOA breaks down.

1.9 Conical Intersections

The case depicted in Figure 1.4 is typical, where the adiabatic states display an *avoided crossing* along a particular coordinate: they become close in energy, but not exactly degenerate. Indeed, in systems with only one internal nuclear degree of freedom (diatomic molecules), and for electronic states of the same spin and spatial symmetry, this is almost always the case. In larger systems, however, the greater number of degrees of freedom allows for situations where the adiabatic states (as well as the diabatic states) are actually degenerate, these are known as *conical intersections*.

A conical intersection point is a particular geometry $\boldsymbol{R}^{\times}$ at which two adiabatic electronic states are exactly degenerate. The degeneracy is lifted in two independent directions or nuclear displacements, any geometrical distortion in these directions causes the PES to split, creating a generic double cone shape that is the origin for the name (see Figure 1.1b). Conversely, distortions in orthogonal directions do not break the degeneracy, which indicates that $\boldsymbol{R}^{\times}$ is not an isolated point, but is part of a connected $3N-8$-dimensional subspace of geometrical configurations, known as intersection space or seam. The points where the energy of the degenerate states is a local minimum within the intersection space are called minimum energy conical intersections (MECI). The two dimensions that do lift the degeneracy form the branching space or branching plane, and are usually denoted as $\boldsymbol{g}$ (the difference between the gradients of the two states) and $\boldsymbol{h}$ (the derivative or non-adiabatic coupling vector). Since the two adiabatic states are exactly degenerate at $\boldsymbol{R}^{\times}$, they are both eigenstates of the Hamiltonian with the same eigenvalue, and therefore any linear combination of them is also an eigenstate with the same eigenvalue. This means that the two states at the intersection are not uniquely defined, and neither are most state-specific properties or the $\boldsymbol{g}$ and $\boldsymbol{h}$ vectors. The branching plane, however – the space spanned by the two vectors – is uniquely defined and independent of which particular linear combinations of the two states is considered.

Conical intersections can be classified according to the topography of the PES in their vicinity. The most common distinction is between *peaked* and *sloped* intersections. Peaked intersections

are exemplified by the ideal shape displayed in Figure 1.1b: when the PES are represented in the branching plane, the intersection point is manifested as local maximum on the lower surface and a local minimum on the upper surface. This latter fact allows them to act as funnels or sinks from the upper to the lower surface. Sloped intersections, on the other hand, can be visualized by tilting the double cone's vertical axis until parts of the upper surface become lower in energy than the vertex (and parts of the lower surface higher than the vertex), at which point the intersection is neither a local maximum nor a local minimum, and its role as attractor is reduced in comparison to peaked intersections. An interesting property of sloped intersections is that they allow the existence of paths that go from the lower surface to the upper surface while continuously decreasing the potential energy.

1.10 Further Reading

For a deeper and more detailed discussion on the topics presented in this introductory chapter, the reader is referred to books on general computational or quantum chemistry, such as those listed below.

- *Modern Quantum Chemistry*. Attila Szabo and Neil S. Ostlund. McGraw–Hill, New York, 1989.
- *Molecular Electronic-Structure Theory*. Trygve Helgaker, Poul Jørgensen and Jeppe Olsen. Wiley, Chichester, 2000.
- *Essentials of Computational Chemistry: Theories and Models*. Christopher J. Cramer. Wiley, New York, 2002.
- *Multiconfigurational Quantum Chemistry*. Björn O. Roos, Roland Lindh, Per Åke Malmqvist, Valera Veryazov and Per-Olof Widmark. Wiley, Hoboken, 2016.
- *Introduction to Computational Chemistry*. Frank Jensen. 3rd ed. Wiley, Chichester, 2017.

1.11 Acknowledgments

L.G. and S.G. thank B. Curchod and F. Agostini for fruitful discussions. I.F.G. and R.L. thank the Swedish Research Council (Grant 2016-03389) and L.G. and S.G. thank the University of Vienna for funding.

Part I

Quantum Chemistry

2

Time-Dependent Density Functional Theory

Miquel Huix-Rotllant, Nicolas Ferré, and Mario Barbatti

Aix Marseille Univ, CNRS, ICR, Marseille, France

Abstract

Linear-response time-dependent density functional theory (LR-TDDFT) in the adiabatic approximation is one of the most popular methods for calculating excited-state properties of molecular systems. This chapter provides a brief overview of the method. We start presenting the formal foundations underlying it. Then, its practical use for calculating vertical excitations and potential energy surfaces, including conical intersections, is accessed. Finally, the use of LR-TDDFT for excited-state nonadiabatic dynamics is discussed. Diverse variants of the method and alternative ways to predict excited states based on DFT are presented as well. Throughout all these sections, our goal has been to highlight the strengths and weaknesses of the method, to provide a guide for best practice when using LR-TDDFT.

2.1 Introduction

Excited states are complicated. Unlike the electronic density of molecules in their ground state, excited-state electronic structure challenges our chemical intuition. Indeed, while the ground state is relatively isolated from the other states, excited states often share a crowded spectral region with states with multiple diabatic characters, distinct multiplicities, and different charges. It is enough to make a slight displacement in the nuclear coordinates and all these states switch their relative energetic order, creating highly anharmonic adiabatic landscapes. This intrinsic complexity of the excited states can even affect the ground state: the forces stabilizing the nuclei in the excited states often drive the molecules to geometries where the ground state can no longer be described by a single configuration.

Until the middle of the 1990s, describing excited states was a demanding task. The performance of the configuration interaction with single excitations (CIS), the workhorse at that time, was pitiful [1]. Although methods such as the complete active space self-consistent field (CASSCF) [2] and multi-reference configuration interaction (MRCI) [3] would deliver qualitatively correct answers, they were prohibitively expensive and would still show large numerical deviations from the experimental values, even at the Franck–Condon region. Moreover, these methods required advanced expertise for the construction and optimization of active and reference spaces (for a recent review on multi-reference methods, see Ref. [4].)

Quantum Chemistry and Dynamics of Excited States: Methods and Applications,
First Edition. Edited by Leticia González and Roland Lindh.

It was around that time, however, that things started to change for excited-state computational research. Methods like the complete active space perturbation theory to the second order (CASPT2) [5], approximated coupled-cluster (CC2, CC3) [6, 7], and time-dependent density functional theory (TDDFT) [8, 9] were developed and implemented in major quantum-chemistry packages. Their accuracy allowed vertical excitation energies to be acquired as near the experimental results as 0.2 eV, and their usage was as close to a black box methodology as we could desire.

The jewel in this popularization of excited-state calculations was the linear-response (LR) TDDFT. Fast, usually reliable, easy to employ, based on conventional ground-state Kohn–Sham (KS) DFT functionals, and counting on analytical energy gradients: it is no wonder the LR-TDDFT quickly established itself as the new routine methodology to aid the assignment of experimental data and analyze reaction mechanism in photoexcited species [10]. Later, with the help of auxiliary wave functions to estimate non-adiabatic couplings [11–13], LR-TDDFT also became a popular option for non-adiabatic mixed quantum-classical (NA-MQC) dynamics, especially in the trajectory surface hopping framework [14].

LR-TDDFT was not bulletproof though. Soon, it was clear that it could dramatically fail to describe the energy in many cases frequently encountered in photochemistry, such as charge-transfer states [10, 15], strongly correlated excited states [16–18], Rydberg states [19], and open-shell excited states [20, 21]. It could also undergo triplet instabilities [22] and deliver conical intersections with the wrong dimensionality [23]; not to mention the problems inherited from the failure of the KS assumption of DFT, such as in bond dissociations or state crossings, incorrect treatment of weak interactions [24], and strong sensitivity to functionals [25].

In the more than two decades since Casida's seminal paper on LR-TDDFT [8], the field of computational excited-state calculations, either with wave function theory or density functional theory, experienced remarkable progress, as we can attest from the other chapters in this book. Nevertheless, LR-TDDFT still shines as the most popular option for the general chemist to interpret spectroscopic data. Due to its operational simplicity and broad availability, LR-TDDFT tends to be uncritically used, with little knowledge of the pathway from the formally exact TDDFT theory underlying it to the approximated LR-TDDFT performed in practice. There is the real risk of misusing LR-TDDFT, working on qualitatively wrong results [26].

The goal of this chapter is to deliver a quick guide into TDDFT, from the formal aspects to its practical use, with a particular focus on its connection to NA-MQC dynamics. Rather than presenting a collection of particular applications, we intend to review the fundamentals and potential problems, to allow a critical and informed use of the method. The reader will be spared the in-depth theoretical details but will be presented with the core of the theory, with plenty of references to follow up. Moreover, we discuss the main handicaps in the method and ways to overcome them. Thus, we hope to provide a guideline for best practice in TDDFT. Working on this text, the reader we bore in mind was a graduate student taking their course on computational chemistry. They may profit from this chapter as a complementary reading to their lecture notes.

2.2 TDDFT Fundamentals

2.2.1 The Runge–Gross Theorems

TDDFT is a formally exact theory [27–37]. In essence, TDDFT states that the time-dependent (TD) density can be used to completely and unequivocally describe a system under a TD external perturbation. Despite initial criticisms to the exact foundations of TDDFT [38–40], it is a general and

well-established theory, grounded on two theorems initially developed by Runge and Gross in 1984 [27], extending the Hohenberg–Kohn theorems of DFT to problems involving TD external potentials [41]. The currently accepted version of the two TDDFT theorems states:

- **Theorem 1: One-to-one correspondence.** There is a one-to-one correspondence between the v-representable TD density $\rho(\mathbf{r}, t)$ and the TD external potential $v_{ext}(\mathbf{r}, t)$ up to an additive time function $\psi(t)$.

$$\rho(\mathbf{r}, t) \overset{\mathbf{j}(\mathbf{r},t)}{\longleftrightarrow} v_{ext}(\mathbf{r}, t) + \psi(t) \tag{2.1}$$

- **Theorem 2: Variational principle.** The exact TD density can be found by applying the least action principle to the Frenkel–Dirac action defined as:

$$A[\rho] = \int_{t_0}^{t_f} dt \left\langle \Psi[\rho](t) \left| i\hbar \frac{\partial}{\partial t} - \hat{H}(t) \right| \Psi[\rho](t) \right\rangle - i\hbar \langle \Psi[\rho](t_f) | \delta\Psi[\rho](t_f) \rangle \,, \tag{2.2}$$

in which $\Psi[\rho](t)$ is the exact time-dependent wave function functional of the time-dependent density, $\delta\Psi[\rho](t)$ is the first-order differential of the wave function, $\hat{H}(t)$ is the time-dependent Hamiltonian, and t_0 and t_f are the initial and final times. As usual, $\hbar$ is the reduced Planck constant.

The first theorem states that the expectation value O of any operator $\hat{O}$ can be evaluated by the TD density, thus, $O[\rho(\mathbf{r}, t)]$. In principle, the expectation values would also depend on the initial state wave function $O[\rho(\mathbf{r}, t); \Psi(\mathbf{r}, t_0)]$ [42]. However, the Hohenberg–Kohn theorems prove the bijective mapping between the initial density and wave function $\rho(\mathbf{r}, t_0) \rightarrow \Psi(\mathbf{r}, t_0)$. Therefore, the operators can be expanded solely by the knowledge of $\rho(\mathbf{r}, t)$. The proof of this theorem is only valid for Taylor expandable external potentials, densities, and currents evaluated around the initial time, given an initial density $\rho(\mathbf{r}, t_0)$ and $\frac{\partial\rho(\mathbf{r},t_0)}{\partial t}$ [27, 31, 34, 37]. The proof is performed in two steps: (i) the bijective mapping between TD external potentials $v_{ext}(\mathbf{r}, t)$ and TD current densities $\mathbf{j}(\mathbf{r}, t)$ is proved, and, then, (ii) the one-to-one correspondence between TD current densities and TD densities through the continuity equation is proved [27, 28]. In fact, TD current densities can be used in the same way as TD densities as an exact theory to evaluate expectation values, a theory known as TD current density functional theory (TDCDFT) [43].

The second theorem was under debate for a long time [29, 32, 38, 40, 44–47]. The initial proposal of Runge and Gross allowed definition of a TDDFT functional, but induced formal problems like the break of causality between the time-dependent perturbation and the time-dependent response [29, 48]. Several solutions have been proposed for defining a TDDFT functional by using Keldysh contour integrations [29], Liouville space pathways [49], or Berry curvatures [32]. The Vignale form of the action principle is shown in Eq. (2.2) [32]. Some of the exact properties of the exact TDDFT functional and its derivatives are known [50].

Perhaps the most remarkable and unique property of TDDFT functionals is the "spatial ultra-non-locality" problem [51]. This property implies that strict local density approximations (LDA) like in DFT are not allowed in TDDFT [52]. The spatial ultra-non-locality specifically appears in functionals of the TD density, but disappear in functionals of the TD current density [53]. There exists also a "time ultra-non-locality" of TDDFT functionals, commonly known as "memory effects," which states that the time evolution of the density depends on the whole history of the density from the initial time t_0 to the current time t [54]. Several memory corrections to the exchange-correlation have appeared in the literature trying to introduce memory character to the time-independent (adiabatic) exchange correlation functionals [55–57]. Direct construction of the exact TDDFT functional is a difficult task (even more difficult than the exact DFT functional), several procedures have appeared to compute exact functional derivatives employing the

time-dependent optimized effective potential procedures [58] or the time-dependent version of the Sham–Schlüter equation [59]. From these approaches, diverse exact TD exchange and TD correlation potentials and kernels have been derived [16, 18, 60–62].

2.2.2 The Time-Dependent Kohn–Sham Approach

From the least-action principle applied to Theorem 2 of Runge and Gross (with the proper definition of the action including Vignale's correction) [27, 28, 32], one can derive a time-dependent Schrödinger equation of the form

$$i\hbar\frac{\partial|\Psi[\rho](t)\rangle}{\partial t} = \hat{H}(t)|\Psi[\rho](t)\rangle\,, \tag{2.3}$$

in which, by means of Theorem 1, the electronic wave function Ψ is a functional of the time-dependent density. In the spirit of the KS procedure for DFT [63], one can construct a TD non-interacting system that will have the same time evolution as the real interacting TD system, with time-dependent electron density

$$\rho(\mathbf{r},t) = \sum_p f_p|\phi_p(\mathbf{r},t)|^2\,, \tag{2.4}$$

in which f_p is the orbital occupation and $\phi_p(\mathbf{r},t)$ is the time-dependent KS spin orbitals. These orbitals follow the Schrödinger-like set of equations commonly known as time-dependent Kohn–Sham (TDKS) equations,

$$i\hbar\frac{\partial\phi_i(\mathbf{r},t)}{\partial t} = \hat{h}^{KS}[\rho](t)\phi_i(\mathbf{r},t)\,. \tag{2.5}$$

In the TDKS, the TD Kohn–Sham Hamiltonian $\hat{h}^{KS}(t)$ is a functional of the time-dependent density, and it is defined as

$$\hat{h}^{KS}(t) = \hat{t}_s + \hat{v}_{ext}(t) + \hat{v}_H[\rho](t) + \hat{v}_{xc}[\rho](t)\,, \tag{2.6}$$

formed by summing the non-interacting kinetic energy $\hat{t}_s$, the time-dependent external potential $\hat{v}_{ext}(t)$, the time-dependent Hartree potential $\hat{v}_H[\rho](\mathbf{r},t)$, and the time-dependent exchange-correlation (xc) potential $\hat{v}_{xc}[\rho](\mathbf{r},t)$. The non-interacting kinetic energy is defined as

$$\hat{t}_s = -\frac{\hbar^2}{2m_e}\nabla^2\,, \tag{2.7}$$

in which m_e is the electron mass, and the time-dependent Hartree potential is

$$\hat{v}_H[\rho](\mathbf{r},t) = \int d\mathbf{r}'\,\frac{\rho(\mathbf{r}',t)}{|\mathbf{r}-\mathbf{r}'|}\,. \tag{2.8}$$

It is important to note that the time-dependent Hartree potential depends only locally on the instantaneous density at time t, but not on previous densities. This is not the case of the TD xc potential, which depends on all the density history from t_0 to the current time.

TDDFT in the TDKS formulation is formally exact (considering the validity of the KS assumption), with the sole condition of knowing the exact TD xc potential. In practice, it has proven to be an arduous task even to obtain an approximation to it. The most widely used approximation is the so-called adiabatic approximation, which completely neglects time-memory effects. In other words, the adiabatic approximation supposes that the TD xc potential depends only on the instantaneous density (hereafter referred as ρ_t to distinguish it from the set of TD densities at all times named ρ). In addition, we can also suppose that the xc functional of DFT (the t_0 xc functional) is

a good approximation to describe the xc effects at all times. Within these two hypotheses, we can straightforwardly define the adiabatic approximation to the TD xc potential as

$$v_{xc}^{A}[\rho](\mathbf{r},t) = \delta(t)\frac{\delta E_{xc}[\rho_t]}{\delta \rho_t(\mathbf{r})}\,. \tag{2.9}$$

Still, the exact DFT functional is also unknown, and further approximations are required to $E_{xc}[\rho_t]$ for practical applications.

2.2.3 Solutions of Time-Dependent Kohn–Sham Equations

There are two basic ways to solve the TDKS equations (Eq. 2.5). We can employ:

- **Real time (RT) integration.** In this case, the TDKS equations are integrated in the real-time domain [64]. The information on the excited states appears through a population transfer between KS orbitals induced by an external field. Having the time-evolution of the dipole momentum, the excitation spectrum is computed by Fourier transforming it [65]. In RT-TDDFT, the nuclei are usually kept frozen during the time integration. Nevertheless, they can be let to evolve coupled to the electrons [66].
- **Linear response (LR) theory.** In the response theory, the poles in the response function occur when the frequency of an external perturbation is resonant with the eigenvalues of the unperturbed system [67]. When the response function is expanded up to linear terms in the perturbation, the approach is called linear-response theory, and the problem is cast as a generalized eigenvalue equation. LR-TDDFT is the popular method usually referred to when we talk about TDDFT.

Both RT and LR approaches are discussed in more detail in the next two sections.

2.2.3.1 Real-Time TDDFT

The TDKS can be integrated directly, as discussed in Ref. [64]. Alternatively, it can be integrated through an unitary time-evolution operator as discussed below.

The TD Schrödinger equation can be conveniently rewritten in terms of a time-evolution operator $\hat{U}(t,t_0)$,

$$i\hbar\frac{d\hat{U}(t,t_0)}{dt} = \hat{H}(t)\hat{U}(t,t_0)\,. \tag{2.10}$$

The evolution operator can propagate a wave function from t_0 to time t,

$$|\Psi(t)\rangle = \hat{U}(t,t_0)|\Psi(t_0)\rangle\,, \tag{2.11}$$

in which the time-evolution operator is formally defined as

$$\hat{U}(t,t') = e^{-i\hbar\int_{t'}^{t} dt'\hat{H}(t')}\,. \tag{2.12}$$

For a non-interacting KS system, one can similarly write a formal solution to the TDKS equation defined in Eq. (2.5) in terms of a non-interacting evolution operator,

$$\phi_i(\mathbf{r},t) = \hat{U}_s(t,t_0)\phi_i(\mathbf{r},t_0)\,, \tag{2.13}$$

in which the non-interacting time-evolution propagator $\hat{U}_s(t,t_0)$ is defined as

$$\hat{U}_s(t,t_0) = e^{-i\hbar\int_{t'}^{t} dt'\hat{h}_{KS}(t')}\,, \tag{2.14}$$

where the TD KS Hamiltonian is defined in Eq. (2.6). Since the exact and the KS Hamiltonians are unitary, the exponentials of the time-evolution operators are anti-Hermitian, and, therefore, they are unitary operators. In addition, due to the additivity of integrals on intervals, one can write

$$\hat{U}(t,t_0) = \hat{U}(t,t')\hat{U}(t',t_0)\,. \tag{2.15}$$

This property is used to perform numerical constructions of $\hat{U}$, by applying the operator n times in intervals of $\Delta t = (t-t_0)/n$,

$$\phi_i(\mathbf{r},t) = \hat{U}_s^n(\Delta t)\phi_i(\mathbf{r},t_0)\,. \tag{2.16}$$

If Δt is sufficiently small, the time-evolution operator can be written as

$$\hat{U}_s(\Delta t) = e^{-i\hbar\hat{h}_{KS}(\Delta t)}\,. \tag{2.17}$$

Numerical approximations of the exponential are then straightforward. For example, the linear term of the Taylor expansion of the operator is

$$\hat{U}_s(\Delta t) \approx \hat{1} - i\hbar\hat{h}_{KS}(\Delta t)\,. \tag{2.18}$$

Using Eq. (2.18) in Eq. (2.16), one can construct the whole evolution of the KS orbitals, which are then used to construct the TD density (Eq. 2.4), from which all properties of the system can be computed. The real-time (RT) solution of TDKS is applicable to weak and intense TD external potentials, and it has a good scaling with the system size since it requires only the knowledge of occupied molecular orbitals. The RT-TDDFT equations can be used in a varied range of applications [68], such as simulation of spectra [69], coupled electron–nuclear dynamics [66], and quantum control [70].

2.2.3.2 Linear-Response TDDFT

In cases for which the external perturbation is weak, it is advantageous to extract the density information using the perturbative approach in the frequency domain. TDDFT in the frequency domain is conveniently formulated using response theory. The basic quantity to construct is the so-called density–density response function, which relates variations in the exact TD density to variations of the TD external potential. Response theory is a powerful technique [71], which allows various types of information to be extracted from the response of the density, not only in the linear regime, but also higher order responses [72], or using special classes of excitation operators such as spin–flip [73] or spin-adapted excitation operators for open-shell systems [74–76]. Here, we discuss the most common form of linear-response TDDFT equations first derived by Casida in 1995 [8], and alternatively derived by Petersilka et al. [77] a year later. An alternative perturbative solution of the LR-TDDFT equations, known as the Sternheimer equation, is another way of solving the linear response of the density [51, 78, 79].

The first-order response (linear response) can be written as

$$\delta\rho(\mathbf{r},t) = \int_{-\infty}^{t} dt' \int d\mathbf{r}'\,\chi(\mathbf{r},\mathbf{r}',t-t')\delta v_{ext}(\mathbf{r}',t')\,, \tag{2.19}$$

in which $\chi(t-t')$ is the response function, and the first-order changes in density and external potential are respectively defined as $\delta\rho(t) = \rho(t) - \rho(t_0)$, $\delta v_{ext}(t) = v_{ext}(t) - v_{ext}(t_0)$. Equivalently, one can write the response function of the non-interacting system as

$$\delta\rho(\mathbf{r},t) = \int_{-\infty}^{t} dt' \int d\mathbf{r}'\,\chi_s(\mathbf{r},\mathbf{r}',t-t')\delta v_s(\mathbf{r}',t')\,, \tag{2.20}$$

in which $\hat{\chi}_s(t-t')$ is the response function based on the wave function for the non-interacting reference state (hereafter referred to as the non-interacting response function) and the first-order change in the TDKS potential is written as $\delta v_s(t) = \delta v_{ext}(t) + \delta v_H(t) + \delta v_{xc}(t)$. Applying the Fourier transform to Eqs. 2.19 and 2.20, one arrives at the so-called Lehmann or sum-over-state representation of the response functions. For the response function based on the wave function of the interacting system (hereafter referred to as the interacting response function), it takes the form (hereafter we omit the space integrations and variables for simplicity)

$$\hat{\chi}(\omega) = \lim_{\eta\to 0^+} \sum_I \frac{\langle\Psi^0|\hat{\rho}|\Psi^I\rangle\langle\Psi^I|\hat{\rho}|\Psi^0\rangle}{\omega - \Omega_I + i\eta} - \frac{\langle\Psi^0|\hat{\rho}|\Psi^I\rangle\langle\Psi^I|\hat{\rho}|\Psi^0\rangle}{\omega + \Omega_I - i\eta}\,, \tag{2.21}$$

in which Ψ^I is the wave function of state I, Ω_I is the true excitation energy of state I, and $\hat{\rho}$ is the density operator. The non-interacting response function takes the form

$$\hat{\chi}_s(\omega) = \lim_{\eta\to 0^+} \sum_{pq} (f_q - f_p) \frac{\hat{\phi}_p \hat{\phi}_p^* \hat{\phi}_q \hat{\phi}_q^*}{\omega - (\epsilon_q - \epsilon_p) + i\eta}\,, \tag{2.22}$$

in which f_p is the occupation number of KS spin orbital ϕ_p and ϵ_p is the KS eigenvalue. Hereafter, we consider that occupation numbers are 0 and 1 only.

It is important to note that when the frequency is equal to the true excitation energy of the system ($\omega = \Omega_I$), the exact response function has a pole ($\hat{\chi}(\Omega_I) = \pm\infty$), while the same is true for the non-interacting response function when the frequency is equal to the difference of eigenvalues ($\hat{\chi}_s(\epsilon_a - \epsilon_i) = \pm\infty$). These criteria are used to extract excitation energies from the interacting system given the knowledge of the non-interacting response function and the exact TD xc kernel. To see this, one needs to substitute the external potential of Eqs. 2.19 in 2.20, and after Fourier transforming, one arrives to the Dyson-type equation for LR-TDDFT

$$(\hat{1} - \hat{\chi}_s(\omega) f_{Hxc}(\omega))\hat{\chi}(\omega) = \hat{\chi}_s(\omega)\,, \tag{2.23}$$

in which we define the exchange-correlation kernel as $f_{Hxc}(\omega) = \frac{\delta v_H(\omega)}{\delta\rho(\omega)} + \frac{\delta v_{xc}(\omega)}{\delta\rho(\omega)}$. When $\omega = \Omega_I$, the left-hand side of Eq. (2.23) tends to infinite, while the right-hand side is finite. To solve this indetermination, the first term in the left-hand side must be 0 when $\omega = \Omega_I$. This is easily imposed by solving the eigenvalue problem

$$(\hat{1} - \hat{\chi}_s(\omega) f_{Hxc}(\omega))X(\omega) = \omega X(\omega)\,. \tag{2.24}$$

This equation is the celebrated linear-response TDDFT equation (LR-TDDFT), which, by expanding it in the KS spin orbital basis, takes the usual Casida form

$$\begin{bmatrix} \mathbf{A}(\omega) & \mathbf{B}(\omega) \\ -\mathbf{B}^*(\omega) & -\mathbf{A}^*(\omega) \end{bmatrix} \begin{bmatrix} \mathbf{X}(\omega) \\ \mathbf{Y}(\omega) \end{bmatrix} = \omega \begin{bmatrix} \mathbf{X}(\omega) \\ \mathbf{Y}(\omega) \end{bmatrix}\,. \tag{2.25}$$

The matrices $\mathbf{A}(\omega)$ and $\mathbf{B}(\omega)$ are defined as

$$A_{ia,jb}(\omega) = \delta_{ij}\delta_{ab}(\epsilon_a - \epsilon_i) + (ia|\hat{f}_{Hxc}(\omega)|bj) \tag{2.26}$$

$$B_{ia,jb}(\omega) = (ia|\hat{f}_{Hxc}(\omega)|jb) \tag{2.27}$$

and we use the Mulliken notation for two-electron integrals

$$(pq|\hat{f}_{Hxc}(\omega)|rs) = \int dr \int dr' \phi_p^*(\mathbf{r})\phi_q(\mathbf{r}) \left[\frac{1}{|\mathbf{r}-\mathbf{r}'|} + f_{xc}(\mathbf{r},\mathbf{r}',\omega)\right] \phi_r^*(\mathbf{r}')\phi_s(\mathbf{r}')\,. \tag{2.28}$$

(For extensions of these integrals for hybrid and range-separated functionals, respectively check equations 15 and 16 of Ref. [80].)

The LR-TDDFT equation presented here is formally exact to determine the first-order response of the TD density at a true excitation energy of the system,

$$\delta\rho(\mathbf{r},\Omega_I) = \sum_{ia}(X_{ia}(\Omega_I)\phi_a^*(\mathbf{r})\phi_i(\mathbf{r}) + Y_{ia}(\Omega_I)\phi_a(\mathbf{r})\phi_i^*(\mathbf{r}))\,, \tag{2.29}$$

from the sole knowledge of the exact $\hat{v}_{xc}(\omega)$ and an initial KS density. However, approximations to Eq. (2.25) are necessary for practical applications. The main common approximations can be summarized as follow:

- **Adiabatic approximation of the TD xc potential:** Employing the adiabatic approximation of the xc potential (Eq. 2.9) in the LR-TDDFT results in a frequency-independent xc terms $\hat{f}_{Hxc}(\omega) \to \hat{f}^A_{Hxc}$. The main consequence is that LR-TDDFT becomes a linear eigenvalue problem and, unlike the exact LR-TDDFT equations, the number of solutions is the same as the dimensions of **A** and **B** matrices. These matrices have the dimensions of one-electron excitations, thus losing the extra solutions corresponding to double and higher electronic excitations. Another consequence is that the xc effects do not take into account the character of the excited state, which can largely affect the quality of the excitation energies obtained by A-LR-TDDFT. This approximation affects only the excited state calculation. Currently, most LR-TDDFT code implementations employ this approximation by default.
- **Approximation of the E_{xc} functional:** If one employs the adiabatic approximation, one is forced to use an approximated form of the DFT exact E_{xc} functional. This will affect the quality of the excitation energies, especially when long-range interactions are required like in charge-transfer states, Rydberg states, etc. This approximation affects both the ground and the excited state calculation. The choice of functional has to describe correctly the virtual orbital space, which is largely influenced by the exchange and correlation functionals.
- **Tamm–Dancoff approximation (TDA):** In TDA, one decouples excitations from de-excitations in the LR-TDDFT, $\mathbf{B}(\omega) = \mathbf{0}$ in Eq. (2.25). The solution of TDA LR-TDDFT form a Hermitian eigenvalue problem of the form $\mathbf{A}(\omega)\mathbf{X}(\omega) = \omega\mathbf{X}(\omega)$ and less sensitive to instabilities of the ground state. This approximation affects only the excited state calculation, and it can improve the description of the energies at the expense of approximate properties [81].

2.3 Linear-Response TDDFT in Action

2.3.1 Vertical Excitations and Energy Surfaces

Adiabatic LR-TDDFT (hereafter we refer to it simply as LR-TDDFT) is by far the most popular approximation of the exact LR-TDDFT approach. Nevertheless, popularity does not necessarily translate into accuracy and generality. For this reason, we will discuss LR-TDDFT successes and failures in some length, having in mind the question whether it is possible to assess the quality of the calculated excitation energies.

In order to understand the physical meaning of the LR-TDDFT equations (see Eq. 2.25), one can apply the so-called single pole approximation (SPA) to a single electron-hole transition as

$$\omega^{SPA} = \epsilon_a - \epsilon_i + (ia|f^A_{Hxc}|ai)\,. \tag{2.30}$$

As can be seen from this approximation, the excitation energies are mainly represented by two ingredients:

- the KS orbital energy gap, which gives a zero order approximation to a one electron excitation;

- the Hartree-exchange-correlation kernel, which give a first-order correction to the orbital energy difference, taking into account not only the coupling between the electron created in orbital a and the hole in orbital i, but the relaxation that affects that modulate the orbital gap [18].

As previously discussed in this chapter, we emphasize that the adiabatic approximation (i.e., the xc potential instantaneously changes with the density) simply wipes out strong correlation effects (represented in wave function theory by double- and higher-order determinants) from LR-TDDFT spectrum [62]. Moreover, the quality of the reference state (i.e., the choice of the ground state exchange-correlation functional, the choice of the basis set, the choice of a geometry, and many other parameters entering the molecular model) also impacts the quality of the LR-TDDFT excitation energies.

In the following, by quality we mean "how close to a state-of-the-art calculation a LR-TDDFT excitation energy is," using the same structure, the same basis set, the same convergence thresholds, etc. This is usually achieved by a systematically benchmarking LR-TDDFT, using different sets of molecules. The extensive review of these benchmarks by Laurent and Jacquemin [82] is definitively a must-read. Of course, the very large number of species in the xc-functional zoo makes the choice heavily system-dependent, ultimately calling for a somehow arbitrary human-made decision. Ideally, this decision should be based on the benchmarking of LR-TDDFT against highly correlated wave function methods.

2.3.1.1 Vertical Excitations: How Good are They?

As a first note regarding the quality of LR-TDDFT vertical excitation energies, it should be remembered that the quality of a higher energy excited state is frequently degraded. This is because of the wrong representation of virtual orbitals due to the incorrect long-range tails of the xc potentials of DFT, which underestimate the ionization potential and place a too low spectral continuum [19]. This induces a mixture of states above the continuum threshold, making them usually too diffuse. This mixture is even more important when the basis set size is increased. In general, it is advised that long-range corrected functionals for TDDFT are used, or at least consider with care states above the ionization threshold (which can be approximated as $-\epsilon_{HOMO}$). Accordingly, for the rest of this section, we only consider the low-lying excited states below the ionization threshold.

Much of the analysis of the quality of the LR-TDDFT vertical excitations has been done around the Mülheim molecular set [83]. This molecular set is a collection of 14 small organic molecules, for which diverse quantum-chemical methods have been computed. It encompasses over 100 singlet–singlet transitions. Tens of functionals have been benchmarked against the theoretical best estimates (TBE) of this molecular set. Using the data collected by Laurent and Jacquemin [82] from diverse references for over 40 functionals, the mean signed errors (MSE) for singlet–singlet vertical excitations are shown in Figure 2.1 by class of functional. We can learn a few different things from this figure:

- The MSE may be anything between −0.5 and +0.6 eV.
- Pure functionals tend to underestimate the excitation energy.
- There is an overall trend of improvement in the order pure → hybrid → range separated → double hybrid, evoking the concept of a Jacob ladder [84].
- Nevertheless, because the variation of MSE within each class of functional may span more of 0.5 eV, there is no guarantee that the results for the higher class will be better than for the lower class.

We also recall that the results in Figure 2.1 correspond to a very homogeneous class of molecules. There are no guarantees that errors (or even the ordering between classes) would still be similar

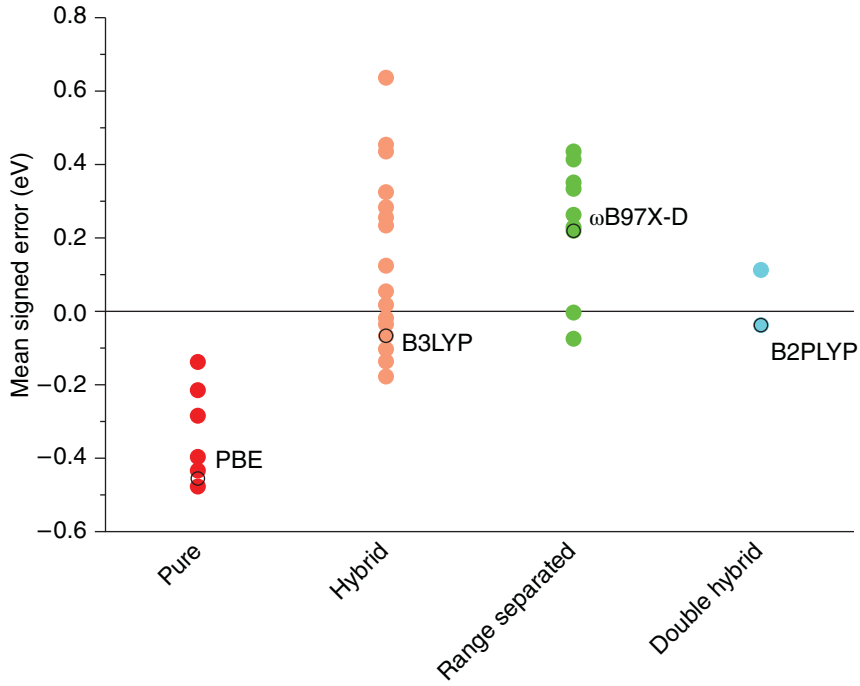


Figure 2.1 Mean signed error (MSE) in the vertical excitation energies (singlet–singlet) in the Mülheim molecular set, for different classes of functionals. Data from Table 2 of Ref. [82]. A few popular functionals are indicated.

to those shown there for different systems (like crystals or metals). In Ref. [85], for instance, the errors for functionals tested on a molecular set of halogenated compounds showed very different trends. Moreover, the description of the excited states should go beyond their energies. For many applications we need to know other properties such as transition dipole moments [86]. It has been reported that for certain cases, a good transition energy does not necessarily correspond to a good description of the electronic density [87].

The data in Figure 2.1 are averages over many different states. Nevertheless, a particular functional may have different performance depending on the type of excitation. Usually, vertical excitations are classified into different types, depending on the nature of the electronic structure modification triggered by the absorption of a photon. Several criteria to diagnostic the states have been developed over the years [88, 89], and a useful quantity to help with the state classification is the Λ parameter [15], available in many quantum chemical programs. Advanced state analysis for large molecular assemblies can also be done with diverse descriptors based on the transition density [90].

The main types of excitations are:

- **Local valence excitations.** When the excited state electronic structure does not differ much from the ground state one, involving, e.g., a typical $\pi \rightarrow \pi^*$ or a $n \rightarrow \pi^*$ single excitation of an organic chromophore or a metal-centered transition in a transition metal complex [91]. Benchmark studies indicate that such excitations can be predicted with an overall good accuracy of about 0.3 eV, if global hybrid xc functionals are used (as we saw, pure functionals underestimate transition energies, while range-separated ones may overestimate them).
- **Rydberg excitations.** They correspond to transitions to hydrogenoid-like and very diffuse excited states. Because of the self-interaction error and the wrong long-range asymptotic

behavior of most common GGA and hybrid functionals, this kind of transition is often characterized by large energy errors ($\gg 0.5$ eV) [92]. Modifications of GGA exchange-enhancement factor can improve the excitation energies [93]. Employing range-separated hybrid functional helps to reduce the discrepancy to about 0.5 eV [94].

- **Charge transfer (CT) excitations.** In a molecular system, they correspond to a transition between two states in which some electron density displaced from the donor subsystem to the acceptor subsystem. (In other words, the electron and the hole do not overlap.) Similarly to Rydberg states, range-separated hybrid functionals are the best choice, usually with an error lower than 0.4 eV [81]. However, if this conclusion applies well to many organic systems, it has to be mitigated when organometallic compounds featuring metal-to-ligand or ligand-to-ligand charge transfer transitions are considered [91]. Among the most recent xc functionals, the double-hybrid ones seem to do an excellent job for both local and charge transfer excitation energies (with a mean error about 0.1 eV) [95] and shall be considered as a promising alternative to commonly used functionals.
- **Core excitations.** These X-ray transitions are characterized by marked electronic structure modifications due to the promotion of an electron from atom-like inner shells to valence orbitals. Core orbitals are extremely compact and will not significantly overlap with valence ones. Accordingly, standard GGA and hybrid xc functionals fail with a very significant (much larger than 10 eV) — but constant — error, which can be pragmatically overcome either through the application of a xc functional-dependent uniform shift, or with a more involved re-parametrization of the xc functional, usually including both short- and long-range Hartree-Fock exchange [96, 97].

We should be cautions about range-separated functionals. These functional have a parameter controlling the separation between the short- and long-range regime [80]; and the value of this parameter has different effects on different types of states. It may, for instance, provide an accurate CT excitation energy at the cost of downgrading the description of local excitations [98]. This reference also shows that the default value adopted in some popular quantum chemistry packages is overestimated.

Of course, the comparison of LR-TDDFT vertical excitations to experimental spectra is not straightforward. Improved models need to take into account other effects, like adiabatic transitions instead of vertical ones and vibronic couplings [99], but also effects of the chromophore surroundings [100], excitonic couplings with other chromophores [89], etc. The comparison with experimental data should also consider that the vertical excitation is in general blue-shifted by about 0.1–0.2 eV in relation to the band maximum [101].

2.3.1.2 Reconstructed Energy Surfaces: How Good are They?

LR-TDDFT computes vertical excitation energies with respect to a reference state calculated at the DFT level, usually the ground state of the chromophore at a given geometry. Accordingly, the excitation energies of all the excited states of interest can be calculated as well. It is then possible to reconstruct the potential energy surfaces of the excited states by systematically changing the nuclear geometric parameters of the molecular system. The broad availability of LR-TDDFT analytical gradients (with respect to nuclear coordinates) [102] and, more recently of analytical Hessian matrices [103], allows location of excited-state stationary energy structures (minima and saddle points), search for minimum energy paths, and to perform excited-state molecular dynamics simulations (see section 2.3.4).

When we systematically explore an adiabatic surface in discrete geometric steps, we usually encounter regions where the character of the adiabatic excited states change between two

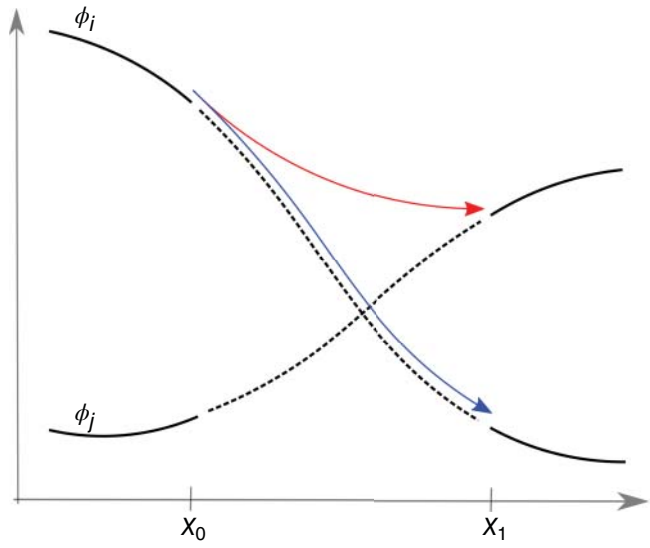


Figure 2.2 Schematic representation of an avoided crossing point between two potential energy curves, occurring between two geometries x_0 and x_1. In x_0 (resp. x_1), the adiabatic lower electronic state is described by the diabatic state Φ_j (resp. Φ_i). The red arrow indicates the "slow" adiabatic pathway, while the blue one indicates the "fast" diabatic one.

consecutive geometries (say, $\mathbf{x}_0$ and $\mathbf{x}_1$, as in Figure 2.2), indicating that an avoided crossing or even a state crossing is located between the two geometries. At such a bifurcation point, two paths become available, with different probabilities: from a Landau–Zener perspective [104], either the system slowly evolves on the same adiabatic state or it ballistically jumps to the other adiabatic state, maintaining its electronic character (Figure 2.2). The characterization of such an event is essential for understanding the molecular photochemistry, as will be highlighted below (section 2.3.2). Therefore, it is clear that the quality of such intersection regions, and more generally, the quality of the LR-TDDFT reconstructed energy surfaces is of tremendous importance to achieve meaningful computational conclusions on photochemical processes.

There are examples in the LR-TDDFT literature reporting bad surface shapes, even if the vertical excitation energy calculated at the Franck–Condon geometry is reasonable. The infamous case of the pentadieniminium cation, the minimum model of retinal (and also known as PSB3), is one such example [105]. It is well known that the ultrafast photoinduced isomerization of PSB3 primarily involves the bond length alternation (BLA) coordinate, eventually turning the isomerizing double bond into a single one [106]. At the same time, close carbon–carbon bonds are also modified to a lower extent. Such BLA alteration manifests itself in the nuclear forces at the Franck–Condon geometry. In the case of LR-TDDFT description of PSB3, looking at the various cases reported in Figure 2.3, the BLA alteration is not satisfactorily reproduced in most of the forces obtained from GGA and hybrid functionals. Although the set of xc-functionals considered in Figure 2.3 is limited, this example clearly shows the extra care required for investigating photochemical processes with LR-TDDFT.

In Figure 2.4, we compare the quality of the potential energy curves obtained from LR-TDDFT against those from CASSCF along the dissociation coordinate of molecular hydrogen. The CASSCF curves show three avoided crossings: (i) between the ground and first excited state at interatomic distances >2 Å, (ii) between the first and second excited state at 0.68 Å, and (iii) between the

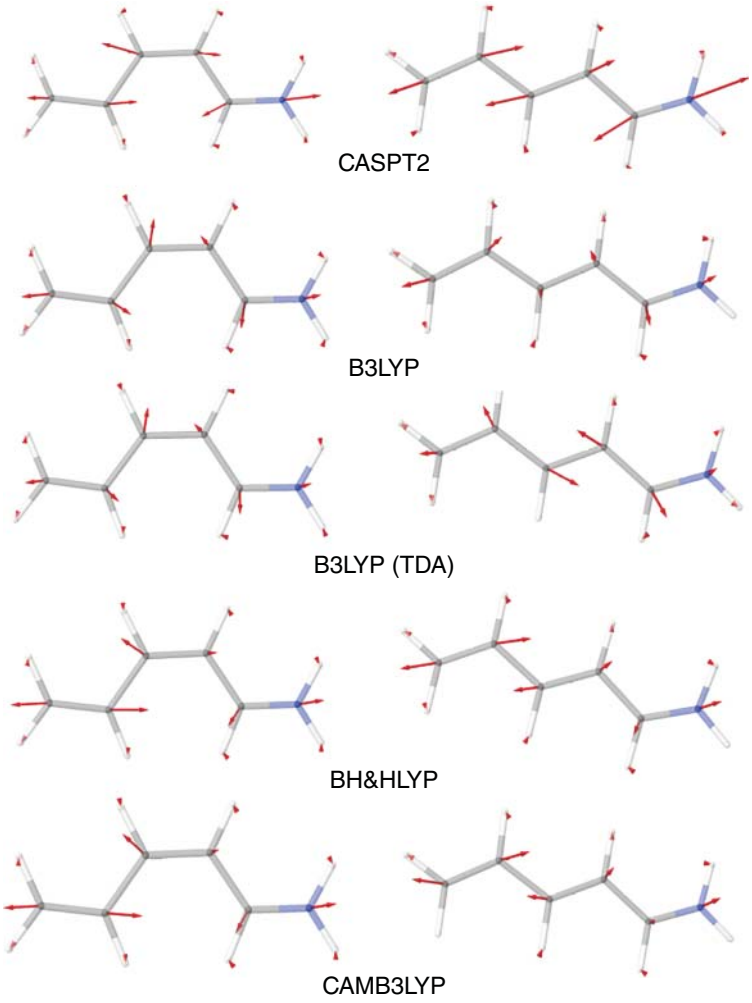


Figure 2.3 PSB3 retinal model LR-TDDFT nuclear forces at the ground state minimum energy geometry. None of the xc functionals is able to capture the correct BLA modification shown by the reference CASPT2 level of theory.

second and the third excited state at 0.95 Å. The first avoided crossing is formed due to the interaction between electronic configurations corresponding to the initial configuration $(\sigma_1)^2(\sigma_1^*)^0$ and the double excitation $(\sigma_1)^0(\sigma_1^*)^2$. The second avoided crossing involves two singly-excited configurations $(\sigma_1)^2(\sigma_1^*)^1(\sigma_2)^0$ and $(\sigma_1)^2(\sigma_1^*)^0(\sigma_2)^1$. Finally, the third avoided crossing is due to the interaction of $(\sigma_1)^2(\sigma_1^*)^0(\sigma_2)^1$ with the double excitation $(\sigma_1)^0(\sigma_1^*)^2$.

None of these crossings is captured by the adiabatic approximation of the TD xc potential. First, A-LR-TDDFT lacks one of the potential energy surfaces corresponding to the doubly-excited configuration $(\sigma_1)^0(\sigma_1^*)^2$. Indeed, the lack of a frequency-dependent xc kernel in the LR-TDDFT equations (Eq. 2.25) results in the lack of doubly-excited configurations. Several consequences can be observed due to this approximation: (i) the potential energy surface of the ground state does not lead to the right dissociation limit of −1.0 Ha. This is, however, a failure that should not be attributed to TDDFT but rather to KS-DFT. (ii) The first excited state is largely overestimated

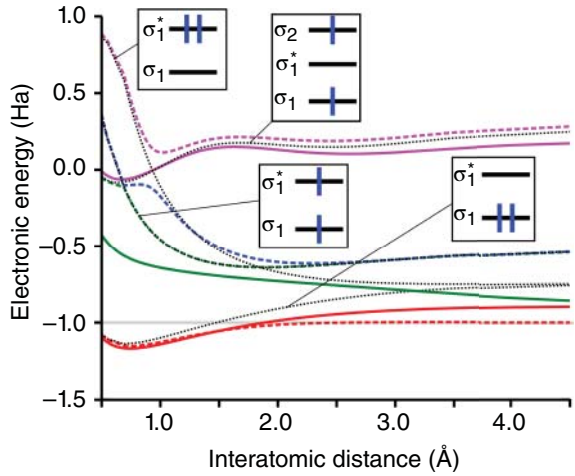


Figure 2.4 Excited states of molecular hydrogen along the dissociation coordinate. The potential energy surfaces of A-LR-TDDFT using B3LYP/6-31G* (solid lines) are compared to the CASSCF(2,3)/6-31G* (dashed lines) and the diabatic surfaces (dotted black lines). The main configuration in each diabatic surface is shown as an offset.

and does not feature the avoided crossing with the second excited state. Such an avoided crossing should in principle be well described in TDDFT, but the wrong description of xc effects in the first excited state underestimates its energy and thus the crossing is absent. (iii) The second excited state at small interatomic distances is absent in LR-TDDFT. This absence is because of the lack of doubly excited configurations in the adiabatic approximation. (iv) The third excited state is represented as a diabatic surface in TDDFT rather than as an adiabatic state as shown in CASSCF. This is indeed another effect of the lack of the double excitation character in A-LR-TDDFT.

Several methods and extensions beyond the A-LR-TDDFT have been proposed to solve the inaccuracies due to a lack of double excitations in LR-TDDFT. Most solutions of the H_2 dissociation imply descriptions beyond the KS assumption from the ground state using fractional occupation numbers, from which the correct avoided crossing of H_2 excited states are extracted. Examples of such theories are time-dependent density-matrix functional theory [107], the constricted variational DFT approach [108], the generalized adiabatic connection in ensemble density-functional theory [109], and restricted ensemble KS (REKS) theory [110]. Another more pragmatic approach is to extend the successful adiabatic xc potential derivatives with frequency-dependence terms. This is the basis of the dressed TDDFT approach of Maitra et al. [16] and generalized by Casida and Huix-Rotllant using many-body perturbation theory [18]. This method features doubly-excited configurations and their interaction with the ground-state configuration, showing an improved surface topology.

2.3.2 Conical Intersections

Conical intersections are state degeneracies characterized by a bi-dimensional linear branching space [111]. They play a central role for ultrafast internal conversion, being the corner-stone of femtochemistry [105]. In 2006, Levine and co-workers [23] raised an important warning: the topology of state intersections described with LR-TDDFT could have the wrong dimensionality. They showed that, although the TDDFT and DFT energies of the state intersection between the first-excited and ground states could be accurate compared to multi-configurational results, the

branching space was anomalously one-dimensional. They tracked down the origin of the problem to the "CIS structure" of the restricted TDDFT (i.e., the exclusive dependence of the method on singly-excited determinants), in which the coupling between the ground state and the excited states is zero by construction. If the closed-shell restriction is lifted by an unrestricted approach, then the branching space becomes bi-dimensional as it should be. Gozem et al. [106] have shown that this same anomalous topology is not exclusive of TDDFT, and it is found in several other methods.

Huix-Rotllant et al. [112] have conducted a systematic investigation of conical intersections computed with wave function theory and density functional methods, including spin-flip DFT and REKS in addition to LR-TDDFT. Benchmarking over eight molecules and twelve conical intersections between S_1 and S_0, they confirmed the findings of Ref. [23] that (1) restricted LR-TDDFT can be used to predict minimum-energy conical intersections but (2) that the dimensionality of the branching space is wrong.

A conical intersection between two states happens when at a certain nuclear geometry $\mathbf{R}_X$ (i) the two diabatic states become degenerated,

$$H_{11}(\mathbf{R}_X) = H_{22}(\mathbf{R}_X), \tag{2.31}$$

and (ii) the diabatic coupling between them is null,

$$H_{12}(\mathbf{R}_X) = 0. \tag{2.32}$$

In these equations, H_{IJ} are the electronic Hamiltonian matrix elements between the diabatic states I and J.

The problems with restricted LR-TDDFT for the description of conical intersections involving the ground state are related to both conditions. First, near the degeneracy (Eq. 2.31), the eigenvalues of the highest occupied and lowest unoccupied KS orbitals tends also to be degenerated, $\varepsilon_{HOMO} \approx \varepsilon_{LUMO}$. Under this situation, DFT may show instabilities, with $\varepsilon_{HOMO} > \varepsilon_{LUMO}$ (sometimes this is referred as a *hole below the Fermi level*) [22]. These instabilities lead to imaginary excitation energies in the LR-TDDFT routine, which, in practical terms, means that the calculations tend to break down near the degeneracy.

Second, LR-TDDFT only calculates excitation energies. They are summed to the DFT ground-state energy to deliver the excited-state energies. Therefore, the Hamiltonian coupling matrix element between the ground and first excited states is always zero, independently of the geometry and not only at $\mathbf{R}_X$ as required by Eq. (2.32). In a multi-reference method, any small nuclear distortion $\delta\mathbf{R}$ along the direction $\partial H_{12}/\partial\mathbf{R}$ lifts the degeneracy because $H_{12}(\mathbf{R}_X + \delta\mathbf{R}) \neq 0$. Nevertheless, in LR-TDDFT, $H_{12}(\mathbf{R}_X + \delta\mathbf{R})$ is still null, causing the wrong dimensionality of the branching space [23]. Tapavicza et al. [113] showed that the first problem in LR-TDDFT can be alleviated by using TDA. According to their analysis, the breaking down of TDDFT near the intersection, where the excitation energy tends to become imaginary, does not occur when TDA is assumed (although the excitation energy can still be negative). TDA alone, however, does not fix the dimensionality problem [114].

The right dimensionality of the intersection seam can be recovered if the excited state is represented by an auxiliary wave function as demonstrated by Li et al. [114] at TDA-TDDFT level. (The construction of these auxiliary wave functions is discussed in Section 2.3.3.1.) In their approach, denominated configuration interaction-corrected Tamm–Dancoff approximation (CIC-TDA), an auxiliary wave function composed of the most relevant singly-excited determinant is used to compute the Hamiltonian matrix element with the closed-shell determinant. This matrix element is, then, used to expand the **A** matrix in TDA to include the closed-shell determinant in the basis. Finally, this new **A** matrix is diagonalized to get the corrected CIC-TDA energies.

Restricted LR-TDDFT fails to describe the topology of the S_1/S_0 intersection due to the lack of coupling between the ground and excited states. This problem does not occur if both the ground and the excited states are calculated as excitations of the same reference. This is exactly what happens in spin-flip TDDFT, which can properly describe the intersection seam with the ground state [115]. For the same reason, the description of conical intersections between excited states is expected to have the correct dimensionality, too.

2.3.3 Coupling Terms and Auxiliary Wave Functions

2.3.3.1 The Casida Ansatz

Neither DFT nor TDDFT produces wave functions by the end of the calculations. Nevertheless, for several types of analysis, we may need specific information about them. This is the case, for instance, of the computation of non-adiabatic couplings between two adiabatic electronic states I and J,

$$\mathbf{d}_{IJ} = \langle \Psi^I | \, \boldsymbol{\nabla}_{\mathbf{R}} \Psi^J \rangle. \tag{2.33}$$

In this equation, the $\boldsymbol{\nabla}_{\mathbf{R}}$ operator indicates the derivatives over the nuclear coordinates $\mathbf{R}$.

Casida [8] showed that the assignment of the excited states calculated with LR-TDDFT can be done using the linear-response coefficients $(\mathbf{X}+\mathbf{Y})^I$ from Eq. (2.25) to build the auxiliary wave function

$$|\Psi^I\rangle = N_I^{-1/2} \sum_{ia} (X+Y)^I_{ia} |\Theta^a_i\rangle, \tag{2.34}$$

where $N_I = \langle (\mathbf{X}+\mathbf{Y}) | (\mathbf{X}+\mathbf{Y}) \rangle$ is the normalization factor. For a functional without Hartree–Fock exchange, $N_I = 1$. Note that in terms of the $\mathbf{F}^I$ coefficients in Casida's original notation, the $(\mathbf{X}+\mathbf{Y})^I$ vectors are

$$(X+Y)^I_{ia} = \sqrt{\frac{\varepsilon_a - \varepsilon_i}{\Delta E_I}} F^I_{ia}, \tag{2.35}$$

where ε_k is the KS energy of orbital k and ΔE_I is the excitation energy of state I.

The auxiliary multi-electron wave function (AMEW) given in Eq. (2.34), which has the simple CIS format, has been often used for computations of a variety of properties, including not only state assignment as originally proposed, but also:

- time-derivative non-adiabatic couplings [11–13];
- non-adiabatic coupling vectors [116];
- spin–orbit couplings [117–119];
- Dyson orbitals [120, 121];
- transition dipole moments [116, 122].

The general theory for using AMEW Ansatz to calculate transition properties has been developed by Tavernelli and co-authors [116, 123].

The use of AMEW as a guess wave function has extrapolated the domains of LR-TDDFT and has been explored in other LR methods like coupled cluster and algebraic diagrammatic construction [26]. Because the calculation of non-adiabatic couplings using such auxiliary functions has been crucial to the development of methods for non-adiabatic dynamics based on LR-TDDFT, we will look at this point in more detail in the next subsection.

Going beyond the use of AMEW, the formal development and implementation of couplings for LR-TDDFT is an active field. Starting from the seminal works of Chernyak and Mukamel [124] and

Baer [125], new methods for couplings between ground and the first excited state [126, 127] as well as between excited states [128–130] have been proposed more recently.

2.3.3.2 Time-Derivative Non-Adiabatic Couplings

In NA-MQC dynamics, the non-adiabatic information at time t enters in the formalism through the time-derivative of the electronic wave function [14],

$$\sigma_{IJ}(t) = \left\langle \Psi^I \left| \frac{\partial \Psi^J}{\partial t} \right. \right\rangle, \quad I \neq J. \tag{2.36}$$

When non-adiabatic coupling vectors $\mathbf{d}_{IJ}$ (Eq. 2.33) are available, the chain rule allows this time-derivative coupling term to be rewritten as $\sigma_{IJ}(t) = \mathbf{v}(t) \cdot \mathbf{d}_{IJ}(t)$, where $\mathbf{v}$ is the nuclear velocity. However, as proposed by Hammes-Schiffer and Tully (HST) [131], σ_{IJ} can also be computed via finite differences

$$\sigma_{IJ}(t) \approx \frac{1}{2\Delta t}\left[S_{IJ}\left(t - \frac{\Delta t}{2}, t + \frac{\Delta t}{2}\right) - S_{IJ}\left(t + \frac{\Delta t}{2}, t - \frac{\Delta t}{2}\right)\right], \tag{2.37}$$

where S_{IJ} are the overlap terms

$$S_{IJ}(t', t) \equiv \langle \Psi^I(t') | \Psi^J(t) \rangle. \tag{2.38}$$

In the previous equation (and in the others in this section), the time dependence is introduced indirectly through the time dependence of the nuclear coordinates $\mathbf{R}(t)$. Thus, $\Psi^I(t)$ should be understood as a short notation for $\Psi^I(\mathbf{R}(t))$

To avoid computing overlaps in between time steps, a linear extrapolation can be used to conveniently rewrite the coupling as [132]

$$\sigma_{IJ}(t) \approx \frac{1}{4\Delta t}[3S_{IJ}(t - \Delta t, t) - 3S_{JI}(t - \Delta t, t) - S_{IJ}(t - 2\Delta t, t - \Delta t) + S_{JI}(t - 2\Delta t, t - \Delta t)]. \tag{2.39}$$

Note that this expression only depends on integers Δt intervals and does not require any information after t. Therefore, its use in dynamics propagation is straightforward.

Now, assuming that the electronic wave function can be written as a linear combination of Slater determinants (just like in the Casida's Ansatz, Eq. 2.34),

$$|\Psi^I\rangle = \sum_{ia} C^I_{ia} |\Theta^a_i\rangle, \tag{2.40}$$

the overlap terms become

$$S_{IJ}(t', t) = \sum_{ijab} C^{I\dagger}_{ia} C^J_{jb} s_{ijab}(t', t), \tag{2.41}$$

where s_{ijab} is the overlap between Slater determinants

$$s_{ijab}(t', t) \equiv \langle \Theta^a_i(t') | \Theta^b_j(t) \rangle. \tag{2.42}$$

(Although we are focusing on singly-excited determinants, this theory is general and applies for any excitation rank.)

If $t = t'$, the s_{ijab} overlap is trivial due to the orthogonality of the Slater determinants. Nevertheless, in dynamics, t and t' are different by Δt, corresponding to slightly different geometries. Löwdin showed long ago [133] that the overlap between determinants built with different spin–orbital bases (ϕ) is given as the determinant of the spin–orbital overlap matrix

$$s_{ijab}(t', t) = \det\{\langle \phi_k(t') | \phi_l(t) \rangle\}, \tag{2.43}$$

where k runs over the spin–orbitals forming $|\Theta^a_i\rangle$ and l over those forming $|\Theta^b_j\rangle$.

As a final step, if the spin–orbitals are formed as a linear combination of atomic orbitals (LCAO), $\tilde{\phi}$, each element in the spin–orbital overlap matrix becomes

$$\langle \phi_k(t')|\phi_l(t)\rangle = \delta_{spin(k),spin(l)} \sum_{\mu\nu} c_{\mu k}(t')c_{\nu l}(t)\langle \tilde{\phi}_\mu(t')|\tilde{\phi}_\nu(t)\rangle. \tag{2.44}$$

$\delta_{spin(k),spin(l)}$ is the Kronecker delta selecting only same-spin terms and $c_{\mu k}$ are the KS orbital coefficients. The overlap between AOs can be obtained as an output of any quantum chemistry program.

Some of the first implementations of the HST (also known as determinant-derivative, DD) algorithm for TDDFT are reported in Refs. [11–13].

The calculation of all spin–orbital overlap matrix elements has a high computational cost. This cost can be reduced if some overlaps terms are neglected when the product of their CI coefficients drops below some threshold or if its excitation rank is too high [13]. If redundancies are eliminated, even larger computational savings are achieved [134].

Following a different approach from HST, Ryabinkin, Nagesh, and Izmaylov (RNI) [135] have shown that, instead of applying the finite difference scheme to Ψ^I as done in Eq. (2.37), the time-derivative coupling σ_{IJ} can be computed by formal differentiation of the AMEW. leading to

$$\sigma_{IJ} = \sum_{ijab} \left[C^{I\dagger}_{ia} \frac{\partial C^J_{jb}}{\partial t} \langle \Theta^a_i | \Theta^b_j \rangle + C^{I\dagger}_{ia} C^J_{jb} \left\langle \Theta^a_i \middle| \frac{\partial \Theta^b_j}{\partial t} \right\rangle \right]. \tag{2.45}$$

Then, in the RNI (or orbital-derivative, OD) algorithm, the finite differences are applied directly to orbitals, after the antisymmetric structure of the determinants is accounted for. Such an algorithm also leads to large computational savings compared to the bare implementation of the HST. (Both HST and NRI algorithms are available in the Newton-X program [136].)

In recent years, the use of wave function overlaps has gained a new function in NA-MQC dynamics beyond the coupling estimate: to detect the so-called trivial or state-unavoided crossings [137]. In situations of weak coupling between states, the non-adiabatic coupling is strongly concentrated in space. It can be so narrow, that it may be missed during the integration of the Newton equations if the time step is too large, leading to wrong predictions of state populations [14]. Several algorithms have been proposed to deal with such narrow coupling functions [138–141], all of them using the overlap elements to keep track of diabatic transformations and correcting the dynamics when needed.

2.3.4 Non-Adiabatic Dynamics

To perform non-adiabatic mixed quantum-classical (NA-MQC) dynamics on-the-fly, we need three basic elements coming from the electronic structure calculations: potential energies of the excited states, energy gradients, and non-adiabatic couplings between the states [14]. Energies and gradients can be routinely computed with any standard TDDFT program. With them, it is already possible to simulate dynamics on a single potential energy surface. Non-adiabatic couplings, however, are not so readily available. The use of AMEW has enabled the computation of couplings during dynamics simulations with LR-TDDFT, especially within the Hammes-Schiffer/Tully approach [131] discussed above. Moreover, implementation of non-adiabatic coupling vectors at TDDFT has also been an area of recent development [116, 127–129].

Usually, the specific NA-MQC dynamics methods (surface hopping (see Chapter 16), Ehrenfest (see Chapter 15), etc.) is blind to the electronic structure method, which produced energies,

gradients, and couplings. Therefore, the use of these three quantities to propagate non-adiabatic dynamics is exactly the same, no matter that they were delivered by LR-TDDFT or, say, ab initio MRCI. For this reason, we will skip the discussion about the dynamics implementation and application. The readers interested in these topics are refereed to chapters 11–20 (see also Refs. [14, 142]). There are a few aspects, however, concerning the dynamics propagation with LR-TDDFT we would like to examine. We list them here:

Conical intersections: Conical intersections in DFT are discussed separately in Section 2.3.2. Here we just summarize that (1) the crossing between exited states may be described with LR-TDDFT and that (2) the dynamics just up the first excited state reaches the crossing point with the ground state may also be fine (although it should be carefully checked). Transitions to the ground state should be avoided with LR-TDDFT.

Momentum correction: In surface hopping, we must adjust the momentum after the molecule hops to another surface to ensure energy conservation [143]. This momentum adjustment, stemming from the Pechukas' forces [144, 145], should be done along the direction of the non-adiabatic coupling vector. If, however, only time-derivative couplings (Eq. 2.36) are computed, this direction is unknown. The usual procedure in such cases is to adjust the momentum along the nuclear momentum direction. Although the total energy will be conserved, there is no theoretical justification for using this direction.

Dissociation and spin contamination: Dissociation is not well described by the Kohn–Sham DFT [35] and naturally this will impact dynamics, if it follows a dissociative coordinate. TDA and unrestricted (U) DFT may alleviate the problem [113]. In the case of U-DFT, we will most likely have to deal with spin contamination.

Charge-transfer states and multiple excitations: All common problems in LR-TDDFT should be expected to occur during the dynamics propagation. Over-stabilization of charge-transfer states [10, 15] and missing multiple excitations [8, 23] are among the most important deficiencies. Charge-transfer problems can be alleviated by using range-separated functionals [146]. There is not much to do about multiple excitations in the context of LR-TDDFT with adiabatic approximation, but to use some auxiliary method to monitor the occasional presence of multiple excitations in the spectral region of interest, and to stop the dynamics if they show up.

Software: At the time of writing, the following are some of the public software enabling NA-MQC with TDDFT:

- DFTBaby - TSH with LR-TD-DFTB [147].
- CPMD - TSH and CT-MQC with LR-TDDFT [12].
- GPAW - MFE with RT-TDKS [148].
- Jade - TSH with LR-TDDFT [149].
- Newton-X - TSH with LR-TDDFT and LR-TD-DFTB [136].
- Octopus - MFE with RT-TDKS [66].
- PYXAID - TSH with RT-TDKS (TD-SDKS) [150].
- QCHEM - TSH with LR-TDDFT [151].
- SHARC - TSH with LR-TDDFT [152].
- Turbomole - TSH with LR-TDDFT (only first excited state) [153].

In this list, TSH stands for trajectory surface hopping [154], MFE for mean-field Ehrenfest [155], and CT-MQC [156] coupled-trajectories mixed quantum-classical dynamics.

2.4 Excited States and Dynamics with TDDFT Variants and Beyond

LR-TDDFT is the most common DFT approach to investigate electronic excited states in molecules and finite molecular assemblies. Nevertheless, it is not the only one [142]. In this section, we compile a brief summary of few complementary and alternative DFT-based methods, which have been used or have the potential to be used for excited-state investigations and non-adiabatic dynamics.

BSE-GW (Bethe–Salpeter equation with GW): The Bethe–Salpeter equation (BSE) applied within the GW approximation has been the most successful method to compute exciton properties in periodic systems typical of condensed-matter physics [157]. More recently, it has begun to be applied to calculate excitation energies in finite systems, being considered a potential way to overcome some of the deficiencies in TDDFT. In practical terms, BSE-GW for finite systems is computed analogously to LR-TDDFT, by solving a similar eigenvalue equation, nevertheless, with different **A** and **B** matrices [158]. These matrices are written in terms of the quasi-particle energies of the many-body GW approach and the screened Coulomb interaction. To the best of our knowledge, BSE-GW has not yet been used in NA-MQC dynamics of molecules mainly due to its computational cost and lack of analytical gradients. It has, however, good potential to be an alternative to LR-TDDFT.

MR-DFT (Multi-reference DFT): Kohn–Sham DFT is based on a single-determinant ground state. There are cases, however, where such a determinant is not flexible enough to describe the state density [159]. For such cases, diverse flavors of MR-DFT have been developed, including (i) semi-empirical MRCI with DFT (e.g., DFT/MRCI by Grimme and Waletzke [160]), (ii) hybrid wave-function DFT (e.g., CAS-srDFT [161, 162]), (iii) multi-configurational DFT (e.g., CAS-DFT [163, 164] and MC-PDFT [165]), and (iv) ensemble DFT (e.g., REKS [166] and GOK-DFT [185]). Among all these methods, the only one that is developed enough to be used in NA-MQC dynamics is REKS. For this reason, it is discussed separately below. Analytical gradients for state-specific ground state have been recently reported for MC-PDFT [167]. Chapter 3 is dedicated to MR-DFT where the interested reader can seek further detail.

REKS (Spin-restricted ensemble-referenced Kohn–Sham): REKS is a rigorous multi-reference DFT approach [166]. It assumes that the ground state density corresponds to that of a weighted average of KS determinants, and variationally builds excited states in small subspaces of electrons and KS orbitals. There have been a series of new developments in REKS, including larger spaces [168], analytical gradients [110], and couplings. Moreover, this method has recently been employed to NA-MQC dynamics [169].

ROKS (Restricted open-shell Kohn–Sham): surface hopping based on ROKS was one of the first on-the-fly NA-MQC methods implemented. Restricted to a single excited state, it was pioneered by Doltsinis and Marx in the early 2000s [170, 171], being applied to investigate the ultrafast dynamics of diverse systems. In their version of ROKS, the excited state is given by the sum of two spin-adapted KS determinants. The KS orbitals for the ground and excited states are determined separately. Differently from the usual Born–Oppenheimer molecular dynamics (BOMD), the Doltsinis–Marx implementation of surface hopping dynamics with ROKS is based on Car–Parrinello molecular dynamics (CPMD) [172].

RT-TDDFT; RT-TDKS (Real-time TDDFT; real-time Kohn–Sham): RT-TDDFT is discussed in Section 2.2.3.1. This electronic dynamics can be directly coupled to the nuclear dynamics in a mean-field Ehrenfest scheme [14] to work on real-space grids [173], plane-waves [174], or atomic orbitals [148]. To adapt RT-TDKS to surface hopping is not straightforward due to the lack of many-electron states required by the fewest switches algorithm [154]. Nonetheless, it has been put forward by Prezhdo et al. [175, 176] as discussed in the entry TD-SDKS below.

sTDDFT (Simplified TDDFT): Bannwarth and Grimme have recently proposed [177] a methodology that strongly reduces the costs of LR-TDDFT, enabling the calculation of molecules with up to 1000 atoms. sTDDFT is based on neglecting some terms in the **A** and **B** matrices in the LR-TDDFT equations, Eq. (2.25); replacing some other terms in these matrices by damped Coulomb interactions; and truncating the single-excitation reference space.

SF-TDDFT (Spin-flip TDDFT): In SF-TDDFT, electronic configurations from excitations operators with $\Delta M_s = \pm 1$ are allowed. Thus, SF-TDDFT applied to an open-shell triplet reference $|ia\rangle$ will generate the singlet configurations $|i\bar{i}\rangle$ (closed shell; the *bar* indicates a β spin), $|i\bar{a}\rangle$ (single excitation), $|\bar{i}a\rangle$ (another single excitation), and $|a\bar{a}\rangle$ (double excitation). As a result, ground and excited singlet states can be treated on the same footing, enabling, for instance, the correct description of conical intersections and even the description of doubly-excited states. The implementation and testing of SF-TDDFT are discussed in Refs. [73, 178]. Dynamics based on SF-TDDFT is discussed in Ref. [179]. Spin contamination can be reduced in SF-TDDFT by using both high-spin components of the triplet state as done in the mixed reference (MR) SF-TDDFT [186].

TDA (Tamm–Dancoff Approximation): TDA is introduced in Section 2.2.3.2. For dynamics, Tapavicza et al. [113] claimed that TDA-TDFT may improve the dynamics results, by reducing the effects of singlet and triplet instabilities.

TD-DFTB (Time-dependent density functional tight binding): DFTB is a parameterized version of DFT [180]. It assumes that the ground-state reference density can be written as a sum of neutral densities of all atoms. This reference is perturbed and the total energy expanded, giving rise to a hierarchy of models (DFTB1, DFTB2 or SCC-DFTB, DFTB3, …). Moreover, DFTB uses a minimum basis set and neglects three-center integrals. Hamiltonian and orbital overlap terms are tabulated from atomic DFT calculations. Thanks to these approximations, DFTB may be thousands of times faster than DFT [181]. Excited states with DFTB can be computed with LR-TD in the same way as TD-DFT [182]. Non-adiabatic dynamics with TD-DFTB has been reported in Refs. [183, 184].

TD-SDKS (Time-dependent single-determinant Kohn–Sham): Aiming at extreme computational savings to allow dynamics for extended molecular assemblies, Prezhdo's group has pioneered a surface hopping method based on excited states represented by single KS determinants [175, 176]. In this approach, which is a surface-hopping variant of RT-TDKS, the excited-state energies are reduced to KS orbitals' energy gaps, while couplings are obtained for the derivative of KS orbitals. Time-dependent coefficients are propagated for each electron independently. They are then assembled into a product to represent a many-electron time-dependent coefficient needed to feed the surface hopping algorithm. The method is available in the PYXAID program [150].

2.5 Conclusions

Throughout this chapter, we have discussed several aspects concerning the time-dependent density functional theory, from its fundamentals to its applications to modeling of potential energy surfaces and non-adiabatic dynamics.

We have seen that the method commonly known as TDDFT corresponds to solving the time-dependent Kohn–Sham formalism using the linear-response theory and neglecting time-memory effects in the exchange-correlation potential, the so-called adiabatic approximation. This methodological formulation owes its popularity to its high computational efficiency and satisfactory accuracy when dealing with many kinds of molecular systems. Nevertheless, a collection of problems entangled in the adiabatic LR theory (and also even in the ground states KS formalism) imposes diverse limitations on the use of the method. Strong dependence on the functional, lack of multiple excitations in the electronic state description, wrong dimensionality

of certain features of the potential energy surface, among other problems, require that we remain cautious at every step of the investigation.

The accumulated experience with LR-TDDFT in the last two decades has taught us a great deal about what we can or cannot do with the method. A large number of computational benchmarks is also available to help to gauge the accuracy of the method for different kinds of application. Some of the main lessons from this common-sense best practice in LR-TDDFT are discussed in this chapter.

The TDDFT research field is still evolving on several different fronts: new functionals either tailored for specific applications or for general improved accuracy, algorithms for calculating diverse properties like analytical Hessians and couplings, parametrized formulations to reduce computational costs, methods to recover multiple excitation effects and to account for strong electron correlation, all these features have been the focus of intense development. In fact, this developmental effort goes far beyond TDDFT, with many novelties being announced in the more general field of excited states based on DFT. We expect that this chapter may serve as an proper introduction to such prolific field.

Acknowledgments

MB thanks the support of the Excellence Initiative of Aix-Marseille University (A*MIDEX) and the project Equip@Meso (ANR-10-EQPX-29-01), both funded by the French Government "Investissements d'Avenir" program. MB also acknowledges funding the WSPLIT project (ANR-17-CE05-0005-01). MHR acknowledges funding for the BIOMAGNET project (ANR-16CE29-0008-01).

References

1 Foresman, J.B., Head-Gordon, M., Pople, J.A., and Frisch, M.J. (1992). Toward a systematic molecular orbital theory for excited states. *J. Phys. Chem.* 96: 135–149.
2 Roos, B.O. (1987). *The Complete Active Space Self-Consistent Field Method and its Applications in Electronic Structure Calculations*, vol. LXIX, 399–445. Hoboken: John Wiley & Sons, Ltd.
3 Shavitt, I. (1998). The history and evolution of configuration interaction. *Mol. Phys.* 94: 3–17.
4 Lischka, H., Nachtigallova, D., Aquino, A.J.A. et al. (2018). Multireference approaches for excited states of molecules. *Chem. Rev.* 118: 7293–7361.
5 Andersson, K., Malmqvist, P.-Å., and Roos, B.O. (1992). Second-order perturbation theory with a complete active space self-consistent field reference function. *J. Chem. Phys.* 96: 1218–1226.
6 Christiansen, O., Koch, H., and Jo/rgensen, P. (1995). Response functions in the CC3 iterative triple excitation model. *J. Chem. Phys.* 103: 7429–7441.
7 Christiansen, O., Koch, H., and Jorgensen, P. (1995). The 2^{nd}-order approximate couped-cluster singles and doubles model CC2. *Chem. Phys. Lett.* 243: 409–418.
8 Casida, M.E. (1995). *Time-Dependent Density Functional Response Theory for Molecules*, 155–192. Singapore: World Scientific.
9 Eric Stratmann, R., Scuseria, G.E., and Frisch, M.J. (1998). An efficient implementation of time-dependent density-functional theory for the calculation of excitation energies of large molecules. *J. Chem. Phys.* 109: 8218–8224.
10 Dreuw, A. and Head-Gordon, M. (2005). Single-reference ab initio methods for the calculation of excited states of large molecules. *Chem. Rev.* 105: 4009–4037.

11 Mitrić, R., Werner, U., and Bonačić-Koutecký, V. (2008). Nonadiabatic dynamics and simulation of time resolved photoelectron spectra within time-dependent density functional theory: ultrafast photoswitching in benzylideneaniline. *J. Chem. Phys.* 129: 164118.

12 Tapavicza, E., Tavernelli, I., and Rothlisberger, U. (2007). Trajectory surface hopping within linear response time-dependent density-functional theory. *Phys. Rev. Lett.* 98: 023001–023004.

13 Pittner, J., Lischka, H., and Barbatti, M. (2009). Optimization of mixed quantum-classical dynamics: time-derivative coupling terms and selected couplings. *Chem. Phys.* 356: 147–152.

14 Crespo-Otero, R. and Barbatti, M. (2018). Recent advances and perspectives on nonadiabatic mixed quantum-classical dynamics. *Chem. Rev.* 118: 7026–7068.

15 Peach, M.J.G., Benfield, P., Helgaker, T., and Tozer, D.J. (2008). Excitation energies in density functional theory: an evaluation and a diagnostic test. *J. Chem. Phys.* 128: 044118.

16 Maitra, N.T., Zhang, F., Cave, R.J., and Burke, K. (2004). Double excitations within time-dependent density functional theory linear response. *J. Chem. Phys.* 120: 5932–5937.

17 Casida, M.E. (2005). Propagator corrections to adiabatic time-dependent density-functional theory linear response theory. *J. Chem. Phys.* 122: 054111.

18 Huix-Rotllant, M. and Casida, M.E. (2016). *Many-Body Perturbation Theory (MBPT) and Time-Dependent Density-Functional Theory (TD-DFT): MBPT Insights About What Is Missing In, and Corrections To, the TD-DFT Adiabatic Approximation*, 445–476. Springer International Publishing.

19 Casida, M.E., Jamorski, C., Casida, K.C., and Salahub, D.R. (1998). Molecular excitation energies to high-lying bound states from time-dependent density-functional response theory: characterization and correction of the time-dependent local density approximation ionization threshold. *J. Chem. Phys.* 108: 4439–4449.

20 Li, Z. and Liu, W. (2016). Critical assessment of TD-DFT for excited states of open-shell systems: I. Doublet-doublet transitions. *J. Chem. Theory Comput.* 12: 238–260.

21 Li, Z. and Liu, W. (2016). Critical assessment of time-dependent density functional theory for excited states of open-shell systems: II. Doublet-quartet transitions. *J. Chem. Theory Comput.* 12: 2517–2527.

22 Casida, M.E., Gutierrez, F., Guan, J. et al. (2000). Charge-transfer correction for improved time-dependent local density approximation excited-state potential energy curves: analysis within the two-level model with illustration for H_2 and LiH. *J. Chem. Phys.* 113: 7062–7071.

23 Levine, B.G., Ko, C., Quenneville, J., and Martínez, T.J. (2006). Conical intersections and double excitations in time-dependent density functional theory. *Mol. Phys.* 104: 1039–1051.

24 Grimme, S. (2011). Density functional theory with London dispersion corrections. *WIREs Comput. Mol. Sci.* 1: 211–228.

25 Medvedev, M.G., Bushmarinov, I.S., Sun, J. et al. (2017). Density functional theory is straying from the path toward the exact functional. *Science* 355: 49–52.

26 Plasser, F., Crespo-Otero, R., Pederzoli, M. et al. (2014). Surface hopping dynamics with correlated single-reference methods: 9H-adenine as a case study. *J. Chem. Theory Comput.* 10: 1395–1405.

27 Runge, E. and Gross, E.K.U. (1984). Density-functional theory for time-dependent systems. *Phys. Rev. Lett.* 52: 997–1000.

28 Gross, E.K.U. and Kohn, W. (1990). Time-dependent density-functional theory. In: *Density Functional Theory of Many-Fermion Systems*, volume 21 of *Advances in Quantum Chemistry* (ed. P.-O. Löwdin), 255–291. Academic Press.

29 van Leeuwen, R. (1998). Causality and symmetry in time-dependent density-functional theory. *Phys. Rev. Lett.* 80: 1280–1283.

30 Burke, K., Werschnik, J., and Gross, E.K.U. (2005). Time-dependent density functional theory: past, present, and future. *J. Chem. Phys.* 123: 062206.

31 Marques, M.A.L., Ullrich, C.A., Nogueira, F. et al. (eds.) (2006). *Time-dependent Density-Functional Theory*, volume 706 of *Lecture Notes in Physics*. Berlin Heidelberg: Springer-Verlag.

32 Vignale, G. (2008). Real-time resolution of the causality paradox of time-dependent density-functional theory. *Phys. Rev. A* 77: 062511.

33 Casida, M.E. (2009). Time-dependent density-functional theory for molecules and molecular solids. *J. Mol. Struct. THEOCHEM* 914: 3–18.

34 Marques, M.A.L., Maitra, N.T., Nogueira, F.M.S. et al. (eds.) (2012). *Fundamentals of Time-dependent Density-Functional Theory*, volume 837 of *Lecture Notes in Physics*. Berlin Heidelberg: Springer-Verlag.

35 Casida, M.E. and Huix-Rotllant, M. (2012). Progress in time-dependent density-functional theory. *Ann. Rev. Phys. Chem.* 63: 287–323.

36 Ullrich, C.A. (ed.) (2013). *Time-dependent Density-functional Theory: Concepts and Applications*. Oxford Graduate Texts. Oxford Scholarship Online.

37 Maitra, N.T. (2016). Perspective: fundamental aspects of time-dependent density functional theory. *J. Chem. Phys.* 144: 220901.

38 Schirmer, J. and Dreuw, A. (2007). Critique of the foundations of time-dependent density-functional theory. *Phys. Rev. A* 75: 022513.

39 Holas, A., Cinal, M., and March, N.H. (2008). Comment on "Critique of the foundations of time-dependent density-functional theory". *Phys. Rev. A* 78: 016501.

40 Schirmer, J. and Dreuw, A. (2008). Reply to "Comment on 'Critique of the foundations of time-dependent density-functional theory' ". *Phys. Rev. A* 78: 056502.

41 Hohenberg, P. and Kohn, W. (1964). Inhomogeneous electron gas. *Phys. Rev.* 136: B864–B871.

42 Maitra, N.T. and Burke, K. (2001). Demonstration of initial-state dependence in time-dependent density-functional theory. *Phys. Rev. A* 63: 042501.

43 Vignale, G. (2012). Time-dependent current density functional theory. In: *Fundamentals of Time-Dependent Density Functional Theory*, volume 837 of *Lecture Notes in Physics* (eds. M. Marques, N. Maitra, F. Nogueira, et al.). Springer.

44 Maitra, N.T., van Leeuwen, R., and Burke, K. (2008). Comment on "Critique of the foundations of time-dependent density-functional theory". *Phys. Rev. A* 78: 056501.

45 Schirmer, J. (2010). Modifying the variational principle in the action-integral-functional derivation of time-dependent density-functional theory. *Phys. Rev. A* 82: 052510.

46 Vignale, G. (2011). Comment on "Modifying the variational principle in the action-integral-functional derivation of time-dependent density-functional theory". *Phys. Rev. A* 83: 046501.

47 Schirmer, J. (2011). Reply to "Comment on 'Modifying the variational principle in the action-integral-functional derivation of time-dependent density functional theory' ". *ArXiv*, abs/1105.2718v1.

48 Gross, E.K., Dobson, J.F., and Petersilka, M. (1996). Density functional theory of time-dependent phenomena. In: *Density Functional Theory II*, volume 181 of *Top. Curr. Chem* (ed. R.F. Nalewajski). Springer-Verlag.

49 Mukamel, S. (2005). Generalized time-dependent density-functional-theory response functions for spontaneous density fluctuations and nonlinear response: resolving the causality paradox in real time. *Phys. Rev. A* 71: 024503.

50 Burke, K. (2006). Exact conditions. In: *Time-dependent Density-functional Theory*, volume 706 of *Lecture Notes in Physics* (eds. M.A.L. Marques, C.A. Ullrich, F. Nogueira, et al.). Berlin Heidelberg: Springer-Verlag.

51 Gonze, X., Ghosez, P., and Godby, R.W. (1995). Density-polarization functional theory of the response of a periodic insulating solid to an electric field. *Phys. Rev. Lett.* 74: 4035–4038.

52 Vignale, G. and Kohn, W. (1996). Current-dependent exchange-correlation potential for dynamical linear response theory. *Phys. Rev. Lett.* 77: 2037–2040.

53 van Faassen, M., de Boeij, P.L., van Leeuwen, R. et al. (2002). Ultranonlocality in time-dependent current-density-functional theory: application to conjugated polymers. *Phys. Rev. Lett.* 88: 186401.

54 Maitra, N.T., Burke, K., and Woodward, C. (2002). Memory in time-dependent density functional theory. *Phys. Rev. Lett.* 89: 023002.

55 Kurzweil, Y. and Baer, R. (2005). Generic Galilean-invariant exchange-correlation functionals with quantum memory. *Phys. Rev. B* 72: 035106.

56 Orestes, E., Capelle, K., da Silva, A.B.F., and Ullrich, C.A. (2007). Generator coordinate method in time-dependent density-functional theory: memory made simple. *J. Chem. Phys.* 127: 124101.

57 Kurzweil, Y. and Baer, R. (2008). Adapting approximate-memory potentials for time-dependent density functional theory. *Phys. Rev. B* 77: 085121.

58 Hirata, S., Ivanov, S., Grabowski, I., and Bartlett, R.J. (2002). Time-dependent density functional theory employing optimized effective potentials. *J. Chem. Phys.* 116: 6468–6481.

59 van Leeuwen, R. (1996). The Sham-Schlüter equation in time-dependent density-functional theory. *Phys. Rev. Lett.* 76: 3610–3613.

60 Görling, A. (1998). Exact exchange kernel for time-dependent density-functional theory. *Int. J. Quantum Chem.* 69: 265–277.

61 Huix-Rotllant, M., Ipatov, A., Rubio, A., and Casida, M.E. (2011). Assessment of dressed time-dependent density-functional theory for the low-lying valence states of 28 organic chromophores. *Chem. Phys.* 391: 120–129. Open problems and new solutions in time dependent density functional theory.

62 Elliott, P., Goldson, S., Canahui, C., and Maitra, N.T. (2011). Perspectives on double-excitations in TDDFT. *Chem. Phys.* 391: 110–119.

63 Kohn, W. and Sham, L.J. (1965). Self-consistent equations including exchange and correlation effects. *Phys. Rev.* 140: A1133–A1138.

64 Yabana, K. and Bertsch, G.F. (1996). Time-dependent local-density approximation in real time. *Phys. Rev. B* 54: 4484–4487.

65 Akama, T. and Nakai, H. (2010). Short-time fourier transform analysis of real-time time-dependent hartree-fock and time-dependent density functional theory calculations with gaussian basis functions. *J. Chem. Phys.* 132: 054104.

66 Andrade, X., Alberdi-Rodriguez, J., Strubbe, D.A. et al. (2012). Time-dependent density-functional theory in massively parallel computer architectures: the octopus project. *J. Phys.: Cond. Matt.* 24: 233202.

67 Olsen, J. and Jørgensen, P. (1995). *Time-Dependent Response Theory with Applications to Self-Consistent Field and Multiconfigurational Self-Consistent Field Wave Functions*, Chapter 13, pages, 857–990. Advanced Series in Physical Chemistry: World Scientific Publishing Company.

68 Marques, M.A.L. and Gross, E.K.U. (2004). Time-dependent density functional theory. *Ann. Rev. Phys. Chem.* 55: 427–455.

69 Marques, M.A.L., Castro, A., Bertsch, G.F., and Rubio, A. (2003). Octopus: a first-principles tool for excited electronion dynamics. *Comput. Phys. Commun.* 151: 60–78.

70 Castro, A., Werschnik, J., and Gross, E.K.U. (2012). Controlling the dynamics of many-electron systems from first principles: a combination of optimal control and time-dependent density-functional theory. *Phys. Rev. Lett.* 109: 153603.

71 Olsen, J. and Jørgensen, P. (1985). Linear and nonlinear response functions for an exact state and for an MCSCF state. *J. Chem. Phys.* 82 (7): 3235–3264.

72 Salek, P., Vahtras, O., Helgaker, T., and Agren, H. (2002). Density-functional theory of linear and nonlinear time-dependent molecular properties. *J. Chem. Phys.* 117: 9630–9645.

73 Shao, Y., Head-Gordon, M., and Krylov, A.I. (2003). The spin-flip approach within time-dependent density functional theory: theory and applications to diradicals. *J. Chem. Phys.* 118: 4807–4818.

74 Li, Z. and Liu, W. (2010). Spin-adapted open-shell random phase approximation and time-dependent density functional theory. I. theory. *J. Chem. Phys.* 133: 064106.

75 Li, Z., Liu, W., Zhang, Y., and Suo, B. (2011). Spin-adapted open-shell time-dependent density functional theory. II. Theory and pilot application. *J. Chem. Phys.* 134: 134101.

76 Li, Z. and Liu, W. (2011). Spin-adapted open-shell time-dependent density functional theory. III. An even better and simpler formulation. *J. Chem. Phys.* 135: 194106.

77 Petersilka, M., Gossmann, U.J., and Gross, E.K.U. (1996). Excitation energies from time-dependent density-functional theory. *Phys. Rev. Lett.* 76: 1212–1215.

78 Sternheimer, R. (1951). On nuclear quadrupole moments. *Phys. Rev.* 84: 244–253.

79 Baroni, S., Giannozzi, P., and Testa, A. (1987). Green's-function approach to linear response in solids. *Phys. Rev. Lett.* 58: 1861–1864.

80 Tawada, Y., Tsuneda, T., Yanagisawa, S. et al. (2004). A long-range-corrected time-dependent density functional theory. *J. Chem. Phys.* 120: 8425–8433.

81 Peach, M.J.G. and Tozer, D.J. (2012). Overcoming low orbital overlap and triplet instability problems in TDDFT. *J. Phys. Chem. A* 116: 9783–9789.

82 Laurent, A.D. and Jacquemin, D. (2013). TD-DFT benchmarks: a review. *Int. J. Quantum Chem.* 113: 2019–2039.

83 Silva-Junior, M.R., Schreiber, M., Sauer, S.P.A., and Thiel, W. (2008). Benchmarks for electronically excited states: time-dependent density functional theory and density functional theory based multireference configuration interaction. *J. Chem. Phys.* 129: 104103.

84 Perdew, J.P., Ruzsinszky, A., Constantin, L.A. et al. (2009). Some fundamental issues in ground-state density functional theory: a guide for the perplexed. *J. Chem. Theory Comput.* 5: 902–908.

85 Stojanović, L., Alyoubi, A.O., Aziz, S.G. et al. (2016). UV excitations of halons. *J. Chem. Phys.* 145: 184306.

86 Robinson, D. (2018). Comparison of the transition dipole moments calculated by TDDFT with high level wave function theory. *J. Chem. Theory Comput.* 14: 5303–5309

87 Plasser, F. and González, L. (2016). Communication: unambiguous comparison of many-electron wave functions through their overlaps. *J. Chem. Phys.* 145: 021103.

88 Etienne, T. (2015). Transition matrices and orbitals from reduced density matrix theory. *J. Chem. Phys.* 142: 244103.

89 Mewes, S.A., Plasser, F., and Dreuw, A. (2015). Communication: Exciton analysis in time-dependent density functional theory: how functionals shape excited-state characters. *J. Chem. Phys.* 143: 171101.

90 Plasser, F., Wormit, M., and Dreuw, A. (2014). New tools for the systematic analysis and visualization of electronic excitations. I. formalism. *J. Chem. Phys.* 141: 024106.

91 Latouche, C., Skouteris, D., Palazzetti, F., and Barone, V. (2015). TD-DFT benchmark on inorganic Pt(II) and Ir(III) complexes. *J. Chem. Theory Comput.* 11: 3281–3289.

92 Caricato, M., Trucks, G.W., Frisch, M.J., and Wiberg, K.B. (2010). Electronic transition energies: a study of the performance of a large range of single reference density functional and wave function methods on valence and rydberg states compared to experiment. *J. Chem. Theory Comput.* 6: 370–383.

93 Li, S.L. and Truhlar, D.G. (2015). Improving Rydberg excitations within time-dependent density functional theory with generalized gradient approximations: the exchange-enhancement-for-large-gradient scheme. *J. Chem. Theory Comput.* 11: 3123–3130.

94 Tsuneda, T. and Hirao, K. (2014). Long-range correction for density functional theory. *WIREs Comput. Mol. Sci.* 4: 375–390.

95 Brémond, É., Savarese, M., Pérez-Jiménez, Á.J. et al. (2017). Speed-up of the excited-state benchmarking: double-hybrid density functionals as test cases. *J. Chem. Theory Comput.* 13: 5539–5551.

96 Besley, N.A. and Asmuruf, F.A. (2010). Time-dependent density functional theory calculations of the spectroscopy of core electrons. *Phys. Chem. Chem. Phys.* 12: 12024.

97 Norman, P. and Dreuw, A. (2018). Simulating X-ray spectroscopies and calculating core-excited states of molecules. *Chem. Rev.* 118: 7208–7248.

98 Sen, K., Crespo-Otero, R., Weingart, O. et al. (2013). Interfacial states in donor-acceptor organic heterojunctions: computational insights into thiophene-oligomer/fullerene junctions. *J. Chem. Theory Comput.* 9: 533–542.

99 Santoro, F. and Jacquemin, D. (2016). Going beyond the vertical approximation with time-dependent density functional theory. *WIREs Comput. Mol. Sci.* 6: 460–486.

100 Guido, C.A., Jacquemin, D., Adamo, C., and Mennucci, B. (2015). Electronic excitations in solution: the interplay between state specific approaches and a time-dependent density functional theory description. *J. Chem. Theory Comput.* 11: 5782–5790.

101 Avila Ferrer, F.J., Cerezo, J., Stendardo, E. et al. (2013). Insights for an accurate comparison of computational data to experimental absorption and emission spectra: beyond the vertical transition approximation. *J. Chem. Theory Comput.* 9: 2072–2082.

102 Furche, F. and Ahlrichs, R. (2002). Adiabatic time-dependent density functional methods for excited state properties. *J. Chem. Phys.* 117: 7433–7447.

103 Liu, J. and Liang, W.Z. (2011). Analytical approach for the excited-state hessian in time-dependent density functional theory: formalism, implementation, and performance. *J. Chem. Phys.* 135: 184111.

104 Desouter-Lecomte, M. and Lorquet, J.C. (1979). Nonadiabatic interactions in unimolecular decay. IV. Transition probability as a function of the massey parameter. *J. Chem. Phys.* 71: 4391–4403.

105 Gozem, S., Luk, H.L., Schapiro, I., and Olivucci, M. (2017). Theory and simulation of the ultrafast double-bond isomerization of biological chromophores. *Chem. Rev.* 117: 13502–13565.

106 Gozem, S., Melaccio, F., Valentini, A. et al. (2014). Shape of multireference, equation-of-motion coupled-cluster, and density functional theory potential energy surfaces at a conical intersection. *J. Chem. Theory Comput.* 10: 3074–3084.

107 Giesbertz, K.J.H., Pernal, K., Gritsenko, O.V., and Baerends, E.J. (2009). Excitation energies with time-dependent density matrix functional theory: singlet two-electron systems. *J. Chem. Phys.* 130: 114104.

108 Ziegler, T., Krykunov, M., and Cullen, J. (2012). The implementation of a self-consistent constricted variational density functional theory for the description of excited states. *J. Chem. Phys.* 136 (12): 124107.

109 Franck, O. and Fromager, E. (2014). Generalised adiabatic connection in ensemble density-functional theory for excited states: example of the H_2 molecule. *Mol. Phys.* 112: 1684–1701.

110 Filatov, M., Liu, F., and Martínez, T.J. (2017). Analytical derivatives of the individual state energies in ensemble density functional theory method. I. General formalism. *J. Chem. Phys.* 147: 034113.

111 Atchity, G.J., Xantheas, S.S., and Ruedenberg, K. (1991). Potential-energy surfaces near intersections. *J. Chem. Phys.* 95: 1862–1876.

112 Huix-Rotllant, M., Nikiforov, A., Thiel, W., and Filatov, M. (2016). *Description of Conical Intersections with Density Functional Methods*, 445–476. Springer International Publishing.

113 Tapavicza, E., Tavernelli, I., Rothlisberger, U. et al. (2008). Mixed time-dependent density-functional theory/classical trajectory surface hopping study of oxirane photochemistry. *J. Chem. Phys.* 129: 124108.

114 Li, S.L., Marenich, A.V., Xu, X., and Truhlar, D.G. (2014). Configuration interaction-corrected Tamm-Dancoff approximation: a time-dependent density functional method with the correct dimensionality of conical intersections. *J. Phys. Chem. Lett.* 5: 322–328.

115 Xu, X., Gozem, S., Olivucci, M., and Truhlar, D.G. (2013). Combined self-consistent-field and spin-flip Tamm-Dancoff density functional approach to potential energy surfaces for photochemistry. *J. Phys. Chem. Lett.* 4: 253–258.

116 Tavernelli, I., Curchod, B.F.E., Laktionov, A., and Rothlisberger, U. (2010). Nonadiabatic coupling vectors for excited states within time-dependent density functional theory in the Tamm-Dancoff approximation and beyond. *J. Chem. Phys.* 133: 194104.

117 de Carvalho, F.F., Curchod, B.F.E., Penfold, T.J., and Tavernelli, I. (2014). Derivation of spin-orbit couplings in collinear linear-response TDDFT: a rigorous formulation. *J. Chem. Phys.* 140: 144103.

118 Gao, X., Bai, S., Fazzi, D. et al. (2017). Evaluation of spin-orbit couplings with linear-response time-dependent density functional methods. *J. Chem. Theory Comput.* 13: 515–524.

119 Dinkelbach, F., Kleinschmidt, M., and Marian, C.M. (2017). Assessment of interstate spin-orbit couplings from linear response amplitudes. *J. Chem. Theory Comput.* 13: 749–766.

120 Arbelo-González, W., Crespo-Otero, R., and Barbatti, M. (2016). Steady and time-resolved photoelectron spectra based on nuclear ensembles. *J. Chem. Theory Comput.* 12: 5037–5049.

121 Humeniuk, A., Wohlgemuth, M., Suzuki, T., and Mitrić, R. (2013). Time-resolved photoelectron imaging spectra from non-adiabatic molecular dynamics simulations. *J. Chem. Phys.* 139: 134104.

122 Mitrić, R., Petersen, J., Wohlgemuth, M. et al. (2011). Time-resolved femtosecond photoelectron spectroscopy by field-induced surface hopping. *J. Phys. Chem. A* 115: 3755–3765.

123 Curchod, B.F.E., Rothlisberger, U., and Tavernelli, I. (2013). Trajectory-based nonadiabatic dynamics with time-dependent density functional theory. *ChemPhysChem* 14: 1314–1340.

124 Chernyak, V. and Mukamel, S. (2000). Density-matrix representation of nonadiabatic couplings in time-dependent density functional (TDDFT) theories. *J. Chem. Phys.* 112: 3572–3579.

125 Baer, R. (2002). Non-adiabatic couplings by time-dependent density functional theory. *Chem. Phys. Lett.* 364: 75–79.

126 Hu, C., Hirai, H., and Sugino, O. (2007). Nonadiabatic couplings from time-dependent density functional theory: formulation in the Casida formalism and practical scheme within modified linear response. *J. Chem. Phys.* 127: 064103.

127 Send, R. and Furche, F. (2010). First-order nonadiabatic couplings from time-dependent hybrid density functional response theory: consistent formalism, implementation, and performance. *J. Chem. Phys.* 132: 044107.

128 Qi, O., Gregory, D., Bellchambers, F.F., and Subotnik, J.E. (2015). First-order derivative couplings between excited states from adiabatic TDDFT response theory. *J. Chem. Phys.* 142: 064114.

129 Li, Z. and Liu, W. (2014). First-order nonadiabatic coupling matrix elements between excited states: a Lagrangian formulation at the CIS, RPA, TD-HF, and TD-DFT levels. *J. Chem. Phys.* 141: 014110.

130 Li, Z., Suo, B., and Liu, W. (2015). First order nonadiabatic coupling matrix elements between excited states: implementation and application at the TD-DFT and pp-TDA levels. *J. Chem. Phys.* 141: 244105.

131 Hammes-Schiffer, S. and Tully, J.C. (1994). Proton-transfer in solution - molecular-dynamics with quantum transitions. *J. Chem. Phys.* 101: 4657–4667.

132 Plasser, F., Granucci, G., Pittner, J. et al. (2012). Surface hopping dynamics using a locally diabatic formalism: charge transfer in the ethylene dimer cation and excited state dynamics in the 2-pyridone dimer. *J. Chem. Phys.* 137: 22A514–22A513.

133 Löwdin, P.-O. (1955). Quantum theory of many-particle systems. I. Physical interpretations by means of density matrices, natural spin-orbitals, and convergence problems in the method of configurational interaction. *Phys. Rev.* 97: 1474.

134 Plasser, F., Ruckenbauer, M., Mai, S. et al. (2016). Efficient and flexible computation of many-electron wave function overlaps. *J. Chem. Theory Comput.* 12: 1207–1219.

135 Ryabinkin, I.G., Nagesh, J., and Izmaylov, A.F. (2015). Fast numerical evaluation of time-derivative nonadiabatic couplings for mixed quantum-classical methods. *J. Phys. Chem. Lett.* 6: 4200–4203.

136 Barbatti, M., Ruckenbauer, M., Plasser, F. et al. (2014). Newton-X: a surface-hopping program for nonadiabatic molecular dynamics. *WIRE: Comput. Mol. Sci.* 4: 26–33.

137 Wang, L. and Prezhdo, O.V. (2014). A simple solution to the trivial crossing problem in surface hopping. *J. Phys. Chem. Lett.* 5: 713–719.

138 Granucci, G., Persico, M., and Toniolo, A. (2001). Direct semiclassical simulation of photochemical processes with semiempirical wave functions. *J. Chem. Phys.* 114: 10608–10615.

139 Spörkel, L. and Thiel, W. (2016). Adaptive time steps in trajectory surface hopping simulations. *J. Chem. Phys.* 144: 194108.

140 Fernandez-Alberti, S., Roitberg, A.E., Nelson, T., and Tretiak, S. (2012). Identification of unavoided crossings in nonadiabatic photoexcited dynamics involving multiple electronic states in polyatomic conjugated molecules. *J. Chem. Phys.* 137: 014512.

141 Meek, G.A. and Levine, B.G. (2014). Evaluation of the time-derivative coupling for accurate electronic state transition probabilities from numerical simulations. *J. Phys. Chem. Lett.* 5: 2351–2356.

142 Barbatti, M. and Crespo-Otero, R. (2016). Surface hopping dynamics with DFT excited states. *Top. Curr. Chem.* 368: 415–444.

143 Fabiano, E., Keal, T.W., and Thiel, W. (2008). Implementation of surface hopping molecular dynamics using semiempirical methods. *Chem. Phys.* 349: 334–347.

144 Pechukas, P. (1969). Time-dependent semiclassical scattering theory. II. Atomic collisions. *Phys. Rev.* 181: 174–185.

145 Herman, M.F. and Kluk, E. (1984). A semiclasical justification for the use of non-spreading wave packets in dynamics calculations. *Chem. Phys.* 91: 27–34.

146 Dreuw, A., Weisman, J.L., and Head-Gordon, M. (2003). Long-range charge-transfer excited states in time-dependent density functional theory require non-local exchange. *J. Chem. Phys.* 119: 2943–2946.

147 Humeniuk, A. and Mitrić, R. (2017). DFTBaby: a software package for non-adiabatic molecular dynamics simulations based on long-range corrected tight-binding TD-DFT(B). *Comput. Phys. Commun.* 221: 174–202.

148 Ojanperä, A., Havu, V., Lehtovaara, L., and Puska, M. (2012). Nonadiabatic Ehrenfest molecular dynamics within the projector augmented-wave method. *J. Chem. Phys.* 136: 144103.

149 Likai, D. and Lan, Z. (2015). An on-the-fly surface-hopping program JADE for nonadiabatic molecular dynamics of polyatomic systems: implementation and applications. *J. Chem. Theory Comput.* 11: 1360–1374.

150 Akimov, A.V. and Prezhdo, O.V. (2013). The PYXAID program for non-adiabatic molecular dynamics in condensed matter systems. *J. Chem. Theory Comput.* 9: 4959–4972.

151 Shao, Y., Gan, Z., Epifanovsky, E. et al. (2015). Advances in molecular quantum chemistry contained in the Q-chem 4 program package. *Mol. Phys.* 113: 184–215.

152 Richter, M., Marquetand, P., González-Vázquez, J. et al. (2011). SHARC: ab initio molecular dynamics with surface hopping in the adiabatic representation including arbitrary couplings. *J. Chem. Theory Comput.* 7: 1253–1258.

153 Tapavicza, E., Meyer, A.M., and Furche, F. (2011). Unraveling the details of vitamin D photosynthesis by non-adiabatic molecular dynamics simulations. *Phys. Chem. Chem. Phys.* 13: 20986–20998.

154 Tully, J.C. (1990). Molecular-dynamics with electronic-transitions. *J. Chem. Phys.* 93: 1061–1071.

155 Tully, J.C. (1998). Mixed quantum-classical dynamics. *Farad. Disc.* 110: 407–419.

156 Min, S.K., Agostini, F., Tavernelli, I., and Gross, E.K.U. (2017). Ab initio nonadiabatic dynamics with coupled trajectories: a rigorous approach to quantum (de)coherence. *J. Phys. Chem. Lett.* 8: 3048–3055.

157 Rebolini, E., Toulouse, J., and Savin, A. (2013). *Electronic Excitation Energies of Molecular Systems from the Bethe-Salpeter Equation: Example of the H_2 Molecule*, 367–389. CRC Press.

158 Krause, K. and Klopper, W. (2017). Implementation of the Bethe-Salpeter equation in the turbomole program. *J. Comput. Chem.* 38: 383–388.

159 Cremer, D. (2001). Density functional theory: coverage of dynamic and non-dynamic electron correlation effects. *Mol. Phys.* 99: 1899–1940.

160 Grimme, S. and Waletzke, M. (1999). A combination of Kohn-Sham density functional theory and multi-reference configuration interaction methods. *J. Chem. Phys.* 111: 5645–5655.

161 Fromager, E. (2015). On the exact formulation of multi-configuration density-functional theory: electron density versus orbitals occupation. *Mol. Phys.* 113: 419–434.

162 Hedegård, E.D., Toulouse, J., and Jensen, H.J.A. (2018). Multiconfigurational short-range density-functional theory for open-shell systems. *J. Chem. Phys.* 148: 214103.

163 Gräfenstein, J. and Cremer, D. (2005). Development of a CAS-DFT method covering non-dynamical and dynamical electron correlation in a balanced way. *Mol. Phys.* 103: 279–308.

164 Gusarov, S., Malmqvist, P.-Å., Lindh, R., and Roos, B.O. (2004). Correlation potentials for a multiconfigurational-based density functional theory with exact exchange. *Theor. Chem. Acc.* 112: 84–94.

165 Manni, G.L., Carlson, R.K., Luo, S. et al. (2014). Multiconfiguration pair-density functional theory. *J. Chem. Theory Comput.* 10: 3669–3680.

166 Filatov, M. (2014). Spin-restricted ensemble-referenced Kohn-Sham method: basic principles and application to strongly correlated ground and excited states of molecules. *WIRE: Comput. Mol. Sci.* 5: 146–167.

167 Sand, A.M., Hoyer, C.E., Sharkas, K. et al. (2018). Analytic gradients for complete active space pair-density functional theory. *J. Chem. Theory Comput.* 14: 126–138.

168 Filatov, M., Martínez, T.J., and Kim, K.S. (2017). Description of ground and excited electronic states by ensemble density functional method with extended active space. *J. Chem. Phys.* 147: 064104.

169 Filatov, M., Min, S.K., and Kim, K.S. (2018). Direct nonadiabatic dynamics by mixed quantum-classical formalism connected with ensemble density functional theory method: application to trans-penta-2,4-dieniminium cation. *J. Chem. Theory Comput.* 14: 4499–4512.

170 Doltsinis, N.L. (2002). *Nonadiabatic Dynamics: Mean-field and Surface Hopping*, volume 10 of *NIC Series*, pages, 377–397. Julich: John von Neumann Institute for Computing.

171 Doltsinis, N.L. and Marx, D. (2002). Nonadiabatic Car-Parrinello molecular dynamics. *Phys. Rev. Lett.* 88: 166402.

172 Car, R. and Parrinello, M. (1985). Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.* 55: 2471–2474.

173 Andrade, X., Castro, A., Zueco, D. et al. (2009). Modified Ehrenfest formalism for efficient large-scale ab initio molecular dynamics. *J. Chem. Theory Comput.* 5: 728–742.

174 Sugino, O. and Miyamoto, Y. (1999). Density-functional approach to electron dynamics: stable simulation under a self-consistent field. *Phys. Rev. B* 59: 2579–2586.

175 Craig, C.F., Duncan, W.R., and Prezhdo, O.V. (2005). Trajectory surface hopping in the time-dependent kohn-sham approach for electron-nuclear dynamics. *Phys. Rev. Lett.* 95.

176 Wang, L., Long, R., and Prezhdo, O.V. (2015). Time-domain ab initio modeling of photoinduced dynamics at nanoscale interfaces. *Ann. Rev. Phys. Chem.* 66: 549–579.

177 Bannwarth, C. and Grimme, S. (2014). A simplified time-dependent density functional theory approach for electronic ultraviolet and circular dichroism spectra of very large molecules. *Comput. Theory Chem.* 1040-1041: 45–53.

178 Xu, X., Yang, K.R., and Truhlar, D.G. (2014). Testing noncollinear spin-flip, collinear spin-flip, and conventional time-dependent density functional theory for predicting electronic excitation energies of closed-shell atoms. *J. Chem. Theory Comput.* 10: 2070–2084.

179 Harabuchi, Y., Keipert, K., Zahariev, F. et al. (2014). Dynamics simulations with spin-flip time-dependent density functional theory: photoisomerization and photocyclization mechanisms of cis-stilbene in $\pi\pi^*$ states. *J. Phys. Chem. A* 118: 11987–11998.

180 Gaus, M., Cui, Q., and Elstner, M. (2014). Density functional tight binding: application to organic and biological molecules. *WIRE: Comput. Mol. Sci.* 4: 49–61.

181 Elstner, M. and Seifert, G. (2014). Density functional tight binding. *Phil. Trans. R. Soc. A ou Philos. Trans. R. Soc.* 372: 20120483.

182 Domínguez, A., Aradi, B., Frauenheim, T. et al. (2013). Extensions of the time-dependent density functional based tight-binding approach. *J. Chem. Theory Comput.* 9: 4901–4914.

183 Mitrić, R., Werner, U., Wohlgemuth, M. et al. (2009). Nonadiabatic dynamics within time-dependent density functional tight binding method. *J. Phys. Chem. A* 113: 12700–12705.

184 Stojanović, L., Aziz, S.G., Hilal, R.H. et al. (2017). Nonadiabatic dynamics of cyclopara-phenylenes with TD-DFTB surface hopping. *J. Chem. Theory Comput.* 13: 5846–5860.

185 Deur, K. and Fromager, E. (2019). Ground and excited energy levels can be extracted exactly from a single ensemble density-functional theory calculation. *J. Chem. Phys.* 150 (9): 094106.

186 Horbatenko, Y., Lee, S., Filatov, M., and Choi, C.H. (2019). Performance analysis and optimization of Mixed-Reference Spin-Flip Time-Dependent Density Functional Theory (MRSF-TDDFT) for vertical excitation energies and singlet–triplet energy gaps. *J. Phys. Chem. A* 123 (37): 7991–8000.

3

Multi-Configurational Density Functional Theory: Progress and Challenges

*Erik Donovan Hedegård**

Division of Theoretical Chemistry, Lund University, Kemicentrum P.O. Box 124, SE-221 00 Lund, Sweden

Abstract

Kohn-Sham density functional theory (KS-DFT) in its time-dependent (linear response) formulation has often provided accurate accounts of electronic excitations and excited state processes. The main features of DFT and its linear response variant are a cost efficient and quite accurate description of dynamical correlation, which is a large challenge for methods based on traditional wave function theory. However, DFT also has limitations, e.g., electronic states where several configurations are of importance; a scenario often encountered for electronically excited states. For these cases a correct description requires a genuine multi-configurational ansatz. The high computational cost of multi-configurational methods has, however, been a large hindrance for their general usage. This has led to development of a number of methods combining DFT and multi-configurational wave functions. The goal is to let the multi-configurational wave function include the configurations required for a physically correct description of the electronic state, while DFT efficiently can recover the dynamical correlation. In this chapter we provide an overview of different methods, combining multi-configurational wave functions and DFT. We discuss the main challenges with outset in three different models namely MRCI/DFT, MC-PDFT and MC-srDFT. These three models have been applied to describe excited states of a large number of organic molecules and to a lesser degree also transition metals; hence an preliminary assessment of their performance for chemically different systems can be made.

3.1 Introduction

Kohn–Sham density functional theory (KS-DFT) has in many ways revolutionized the role of computational and theoretical chemistry, also regarding electronically excited states [1–4]. The success of DFT can often be attributed to a reliable representation of the short-range *dynamical correlation*. For excited states, DFT has become particularly popular in its time-dependent formulation, usually denoted TD-DFT. Yet, it is today clear that the approximations employed for the exchange–correlation functionals impose limitations. These limitations become evident for molecular systems where several configurations are of importance. In this case a correct description requires a genuine multi-configurational *ansatz* [5, 6]. The electron-correlation associated with this scenario is denoted *static correlation*. Describing static correlation is a general problem of DFT, both

* Electronic mail: erik.hedegard@teokem.lu.se

Quantum Chemistry and Dynamics of Excited States: Methods and Applications,
First Edition. Edited by Leticia González and Roland Lindh.

for ground and excited states. In this context, it should be mentioned that TD-DFT avoids explicit optimization of the excited state wave function, and hence may work well if the ground-state is not multi-configurational, although a number of conditions still need to be fulfilled. For instance, states corresponding to configurations higher than singly excited with respect to the ground state are not described by common DFT approximations (for developments addressing these issues, see Refs. [7, 8] or Chapter 2). Moreover, excitations of charge-transfer (CT) [9–13] or Rydberg character are known to be drastically underestimated by TD-DFT.

Development of wave function methods to handle multi-configurational systems is an active field of research [6]. One of the main obstacles faced when applying these wave functions is a prohibitively large computation cost. This has led to a large variety of different multi-configurational schemes. The most common schemes define a *complete active space* CAS(m, n) of m electrons in n orbitals in which all configurations (fulfilling additional spin and symmetry constraints) are included. This is usually combined with optimization of orbital parameters in what is denoted a complete-active-space self-consistent field (CASSCF) procedure. Unfortunately, many systems require active space sizes beyond the current limitations to give physically meaningful results. Several groups have focused on lifting the limitations for the size of active space with methods such as the density-matrix renormalization group (DMRG) [14–17], quantum Monte Carlo (QMC) [18–20], restricted-active-space (RAS) [21] or generalized-active-space (GAS) [22–24]. These methods are described in chapters 7, 8 and 6, respectively. Yet, even with extended active orbital spaces, essential parts of the dynamical electron correlation cannot be obtained, except for the smallest systems. Typically, dynamical correlation for such multi-configurational systems is obtained *after* initially obtaining a correct representation of a zeroth-order Hamiltonian (including static correlation). The exact nature of the subsequent steps responsible for recovering dynamical correlation depends on the chosen method, but well-known examples are multi-configurational perturbation theories [25–29] such as complete-active-space second-order perturbation theory (CASPT2) [25, 26] and n-electron valence state perturbation theory (NEVPT2) [29].

A computationally cheap alternative to obtain dynamical correlation with multi-configurational wave functions is to combine them with a DFT functional and thereby exploit the efficiency of DFT, while maintaining a correct description of multi-configurational systems. In turn, the combination can also correct the errors observed in current DFT approximations for excited states. Still, the combination is not straightforward, since static and dynamical correlation are merely phenomenological terms and cannot be strictly separated. A part of the dynamical correlation is therefore always included in the multi-configurational wave function, leading to double-counting of correlation effects if no special precautions are taken in the combination. The double-counting issue has been addressed in different ways but there is currently no consensus on the best strategy. This chapter provides a summary of the main developments in combining DFT and multi-configurational wave functions, including how the different methods address the double-counting issue. The field has been active for many years, and constitutes by now a vast body of literature. Larger overviews have been given recently, e.g., by Ghosh et al. [30] and Lischka et al. [31]. Following Lischka et al. [31], one can roughly divide multi-configurational DFT into four different families, although the boundaries between these families are loose: the different families are listed in Table 3.1. A complete overview of these different families is beyond the scope of this chapter. Instead, we discuss the progress and challenges in the field with attention on a few, selected methods, focusing on methods that have been applied to describe excited states. This more narrow focus will also allow us to scrutinize the mathematical framework

Table 3.1 Overview over different families of multi-configurational wave functions and DFT.

Class	Examples	Excited states	Selected references
Conventional	DFT	State-specific (ΔSCF)/TD-DFT/Real-time	2–4
Semi-empirical	MRCI/DFT	State-specific/State-average	32, 33, 47
On-top pair density	MC-PDFT	State-specific/State-average	34, 35, 48
Range-separated	MC–srDFT	State-specific/State-average/TD-MC-srDFT	36,49–52
Ensemble DFT	REKS	State-specific/State-average	43–46

of the selected methods in more detail. We have chosen to focus on methods that are in active development, namely MRCI/DFT [32, 33], multi-configurational pair-density functional theory (MC-PDFT) [34, 35] and multi-configurational ranged-separated short-range DFT (MC–srDFT) [36, 37]. The chosen methods belong to different families, and have been tested on a number of different systems for electronically excited states. We will in the last section use this to illustrate their performance for vertical excitation energies. The first method we describe (MRCI/DFT) is a semi-empirical method. In these methods, DFT functionals from regular KS theory are added to a multi-configurational wave function. Double counting is minimized by either scaling the density employed in the DFT functional, or by scaling matrix elements of the Hamiltonian. The next two methods (MC-PDFT and MC-srDFT) are more elaborate and avoid double-counting by either re-defining the DFT functionals in terms of a so-called on-top pair density or decompose the electron–electron repulsion into long- and short-range parts. In Table 3.1, it is also indicated that the three methods can utilize different strategies to calculate excited states and the corresponding excitation energies. We comment on this for the individual methods. When discussing the individual methods, references to methods building on similar ideas will be given, but the reference list is not intended to be exhaustive. Moreover, it should also be noted that the class of methods building on ensemble DFT [38–40] in Table 3.1 is left out of this overview. This does not reflect lack of development in this branch of multi-configurational DFT models. In fact, quite the opposite is the case as illustrated by recent works by Pernal and co workers [41] as well as the work by Franck and Fromager [42]. These methods have, however, not yet been throughly benchmarked. Their performance in excited state calculations over a larger set of systems is therefore yet to be seen. A number of important contributions have also been made by Filatov in what is denoted the spin-restricted ensemble-references Kohn–Sham (REKS) method [43–45]. The REKS method is among the few multi-configurational DFT models developed for conical intersections [46]; an area where TD-DFT can be problematic. Our main argument to leave out ensemble DFT is length considerations and that REKS and conical intersections are discussed in more depth in Chapter 2.

This chapter starts with brief overviews of wave function theory (Section 3.2), and density functional theory (Section 3.3). The latter is also treated in the chapter on TD-DFT, but is retained here to make this chapter self-contained. These two first sections will introduce the nomenclature required to discuss the multi-configurational hybrid models in Section 3.4. Further, Section 3.3 will also discuss common approximations applied in DFT (and TD-DFT), and their consequences for calculations on excited states. In Section 3.5, we investigate how the models discussed on Section 3.4 perform for vertical excitation energies, while an outlook is given in Section 3.6.

3.2 Wave Function Theory

We begin with the time-independent, electronic (non- or scalar-relativistic) Hamiltonian, $\hat{H}$, whose eigenstates, $|\Psi\rangle$, are solutions to the time-independent Schrödinger equation

$$\hat{H}|\Psi\rangle = E|\Psi\rangle, \tag{3.1}$$

where $\hat{H}$ is given in Chapter 1 (deviations from the nomenclature given in Chapter 1 will be explicitly stated). For brevity, we introduce a short form of the electron repulsion operator

$$g_{ij} = \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}. \tag{3.2}$$

As discussed in Chapter 1, the parametrization of a multi-configurational wave function is most convenient in a second-quantization formulation [53]. A difference to Chapter 1 is that we will sometimes require explicit reference to spin-indices in the one- and two-particle density matrices

$$\bar{D}_{p\sigma q\sigma} = \langle\Psi|\hat{a}^\dagger_{p\sigma}\hat{a}_{q\sigma}|\Psi\rangle \tag{3.3}$$

$$\bar{\Gamma}_{p\sigma q\sigma r\tau s\tau} = \langle\Psi|\hat{a}^\dagger_{p\sigma}\hat{a}^\dagger_{r\tau}\hat{a}_{s\tau}\hat{a}_{q\sigma}|\Psi\rangle, \tag{3.4}$$

which are related to the matrices given in Chapter 1 as

$$D_{pq} = \langle\Psi|\hat{E}_{pq}|\Psi\rangle = \sum_\sigma \bar{D}_{p\sigma q\sigma} \tag{3.5}$$

$$\Gamma_{pqrs} = \langle\Psi|\hat{e}_{pqrs}|\Psi\rangle = \sum_{\sigma\tau} \bar{\Gamma}_{p\sigma q\sigma r\tau s\tau}. \tag{3.6}$$

The spin-averaged one- and two-electron replacement operators ($\hat{E}_{pq}$ and $\hat{e}_{pqrs}$) are defined in terms of the creation ($\hat{a}^\dagger_{p\sigma}$) and annihilation ($\hat{a}_{q\sigma}$) operators (see Chapter 1 for their definition).

3.3 Kohn–Sham Density Functional Theory

The most commonly employed form of DFT is the KS formulation, and the multi-configurational extensions discussed in the next section also builds on this formulation. In the KS framework, the total energy is described as a functional of the electron density, divided into the terms

$$E_{\mathrm{DFT}}[\rho] = T_S[\rho] + E_{\mathrm{ext}}[\rho] + E_{\mathrm{H}}[\rho] + E_{\mathrm{xc}}[\rho]. \tag{3.7}$$

The individual terms are described in more detail below. The first term in Eq. (3.7) is the kinetic energy, which in KS-DFT is obtained as the expectation value over a single-determinant wave function, $|\Phi\rangle$

$$T_S[\rho] = \sum_i \langle\Phi|\hat{T}_i|\Phi\rangle, \tag{3.8}$$

where $\hat{T}_i$ is the kinetic energy operator. The single-determinant wave function is comprised of orbitals $\{\phi_i(\mathbf{r})\}$ and is an exact wave function for a system of non-interacting electrons; hence $|\Phi\rangle$ is denoted the non-interacting reference system. The connection to the real system is obtained by requiring that the non-interacting reference system has the same electron density

$$\rho(\mathbf{r}) = \sum_i \phi^*_i(\mathbf{r})\phi_i(\mathbf{r}) = \sum_i |\phi(\mathbf{r})|^2, \tag{3.9}$$

as the real system with wave function $|\Psi\rangle$,

$$\rho(\mathbf{r}) = N \int \dots \int \Psi^*(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)\mathrm{d}\sigma_1 \mathrm{d}\mathbf{x}_2 \mathrm{d}\mathbf{x}_3 \cdots \mathrm{d}\mathbf{x}_N, \tag{3.10}$$

where $\mathbf{x}$ denotes spatial ($\mathbf{r}$) and spin variables (σ), respectively.

An obvious difference between the non-interacting and real systems is that the real system experiences electron-repulsion. In the Kohn–Sham scheme the electron-repulsion is decomposed into a classical term, $E_{\mathrm{H}}[\rho]$, and a non-classical part, $E_{\mathrm{xc}}[\rho]$. With this decomposition, all classical electrostatic terms are contained in $E_{\mathrm{ext}}[\rho]$ and $E_{\mathrm{H}}[\rho]$

$$E_{\mathrm{H}}[\rho] = \frac{1}{2} \iint \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 \tag{3.11}$$

$$E_{\mathrm{ext}}[\rho] = \sum_i \langle \Phi | \hat{V}_{\mathrm{ext}}(\mathbf{r}_i) | \Phi \rangle = \int \rho(\mathbf{r})\ \hat{V}_{\mathrm{ext}}(\mathbf{r}) \mathrm{d}\mathbf{r}. \tag{3.12}$$

In absence of other external potentials, $\hat{V}_{\mathrm{ext}}(\mathrm{r}_i)$ is comprised solely of the nuclear–electron attraction term, represented by $\hat{V}_{\mathrm{ne}}(\mathbf{r}_i)$ in Chapter 1. The non-classical parts of the electron–electron interaction (exchange and correlation) are collected in the exchange-correlation functional, $E_{\mathrm{xc}}[\rho]$. The true (exact) kinetic energy, $T[\rho]$, is obtained after adding a correction, $T_C[\rho]$, to the kinetic energy obtained from a single Slater determinant, i.e., $T[\rho] = T_S[\rho] + T_C[\rho]$. Thus, it is apparent by isolating $E_{\mathrm{xc}}[\rho]$ in Eq. (3.7) that $E_{\mathrm{xc}}[\rho]$ formally also contains a kinetic energy contribution in addition to the contributions from exchange and correlation. The expression for $E_{\mathrm{xc}}[\rho]$ is thus

$$E_{\mathrm{xc}}[\rho] = T[\rho] - T_S[\rho] + E_{\mathrm{ee}}[\rho] - E_{\mathrm{H}}[\rho]. \tag{3.13}$$

To combine more elaborate (multi-configurational) wave functions and DFT, it is convenient to formulate DFT in second-quantization. We here follow Refs. [54, 55] and parameterize the Kohn–Sham reference state as

$$|\Phi(\boldsymbol{\kappa})\rangle = e^{-\hat{\kappa}} |\Phi\rangle, \tag{3.14}$$

where the orbital coefficients, $\boldsymbol{\kappa}$, are given within the orbital-rotation operator

$$\hat{\kappa} = \sum_{pq} \kappa_{pq} \hat{E}_{pq} = \sum_{p>q} \kappa_{pq} (\hat{E}_{pq} - \hat{E}_{qp}) \equiv \sum_{p>q} \kappa_{pq} \hat{E}^-_{pq}, \tag{3.15}$$

and $\hat{\kappa}$ is defined as an antisymmetric, real singlet operator (see also Chapter 1). We can now define a charge-density operator for Eq. (3.10) and translate the density functionals in Eqs. (3.11), (3.12) and (3.13) to second-quantization [54, 55]. Although we only require the charge-density for the exact expressions, we here follow Ref. [37] and introduce a more general nomenclature that allows definition of both charge- and spin-densities; the latter is often employed in approximate exchange correlation functionals. We define the density operators

$$\hat{\rho}^x(\mathbf{r}) = \sum_{pq} \Omega_{pq}(\mathbf{r}) \hat{X}_{pq}, \tag{3.16}$$

where $\Omega_{pq}(\mathbf{r}) = \phi^*_p(\mathbf{r})\phi_q(\mathbf{r})$. In this nomenclature, the superscript $x = c$ is used for the regular charge-density operator (in this case $\hat{X}_{pq} \equiv \hat{E}_{pq}$), while $x = s$ denotes the spin-density operator that goes with

$$\hat{X}_{pq} \equiv \hat{T}_{pq} = \hat{a}^\dagger_{p\alpha} \hat{a}_{q\alpha} - \hat{a}^\dagger_{p\beta} \hat{a}_{q\beta}, \tag{3.17}$$

which is required to describe spin-density effects (in a non-relativistic framework). As noted in Ref. [37], for a spin-restricted, multi-configurational model (which encompasses all models discussed here), the charge- and spin-density operators are singlet and triplet operators (both of

$M_S = 0$ type), respectively. The electron charge or spin densities can for any wave function, $|\Psi\rangle$, be obtained as

$$\rho^x(\mathbf{r}) = \langle\Psi|\hat{\rho}^x|\Psi\rangle = \sum_{pq} \Omega_{pq}(\mathbf{r}) D^x_{pq}, \tag{3.18}$$

where D^x_{pq} is the (p,q) element of the one-particle density matrix (cf. Eq. 3.5) or the one-particle spin-density matrix, $D^s_{pq} = \bar{D}_{p\alpha q\alpha} - \bar{D}_{p\beta q\beta}$. We note that it is also possible to define $\rho^\alpha(\mathbf{r})$ and $\rho^\beta(\mathbf{r})$ densities according to

$$\rho^\sigma(\mathbf{r}) = \langle\Psi|\hat{\rho}^\sigma|\Psi\rangle = \sum_{pq} \Omega_{pq}(\mathbf{r})\langle\Psi|\hat{a}^\dagger_{p\sigma}\hat{a}_{q\sigma}|\Psi\rangle = \sum_{pq} \Omega_{pq}(\mathbf{r})\bar{D}_{p\sigma q\sigma}, \tag{3.19}$$

and thus decompose the total electron density in Eq. (3.10) in terms α and β spin-densities [56]

$$\rho(\mathbf{r}) = \sum_\sigma \rho^\sigma(\mathbf{r}) = \rho^\alpha(\mathbf{r}) + \rho^\beta(\mathbf{r}). \tag{3.20}$$

Both $\rho^s(\mathbf{r})$, $\rho^\alpha(\mathbf{r})$ and $\rho^\beta(\mathbf{r})$ can be employed for approximate exchange–correlation functionals. In the conventional KS formalism, Eq. (3.18) is obtained with the KS determinant and the one-particle density matrix is thus obtained as $\mathbf{D}^{\mathrm{SCF}}$, leaving the energy as

$$E_{\mathrm{DFT}}[\rho] = \sum_{pq} h_{pq} D^{\mathrm{SCF}}_{pq} + \frac{1}{2}\sum_{pqrs}(pq|rs) D^{\mathrm{SCF}}_{pq} D^{\mathrm{SCF}}_{rs} - \frac{1}{2}\gamma_{\mathrm{HF}} \sum_{pqrs}(pq|rs) D^{\mathrm{SCF}}_{ps} D^{\mathrm{SCF}}_{rq} + E_{\mathrm{xc}}[\rho], \tag{3.21}$$

with the one- and two-electron integrals defined as in Chapter 1. Although we keep the more elaborate forms in this part, it should be noted that $\mathbf{D}^{\mathrm{SCF}}$ attains a rather simple (diagonal) form and the energy expression thus simplifies as expected from the first-quantized form in Eqs. (3.9) and (3.11)–(3.12).

The energy in Eq. (3.21) can be obtained by solving a set of self-consistent equations, known as the Kohn–Sham equations. The equations are formulated as the eigenvalue problem

$$\hat{f}^{\mathrm{KS}}|\Phi\rangle = \epsilon|\Phi\rangle, \tag{3.22}$$

where ϵ is a diagonal matrix and $\hat{f}^{\mathrm{KS}}$ is the Kohn–Sham operator,

$$\hat{f}^{\mathrm{KS}} = \sum_{pq}\sum_\sigma f^{\mathrm{KS}}_{pq}\hat{a}^\dagger_{p\sigma}\hat{a}_{q\sigma}, \tag{3.23}$$

with matrix elements

$$f^{\mathrm{KS}}_{pq} = h_{pq} - \sum_{rs}\sum_\sigma((pq|rs) - \gamma_{\mathrm{HF}}(pr|sq))\bar{D}^{\mathrm{SCF}}_{r\sigma s\sigma} + \int \hat{v}_{\mathrm{xc}}\Omega_{pq}(\mathbf{r})\mathrm{d}\mathbf{r}. \tag{3.24}$$

We have in Eq. (3.24) again refrained from simplification, and kept $\bar{\mathbf{D}}^{\mathrm{SCF}}$. The potential from the exchange–correlation energy in Eq. (3.24) is defined $\hat{v}_{\mathrm{xc}} = \frac{\delta E_{\mathrm{xc}}}{\delta\rho}$, while the integrals have been defined previously (cf. Chapter 1). We note that the Coulomb integrals are sometimes written

$$(pq|rs) = \iint \frac{\Omega_{pq}(\mathbf{r}_1)\Omega_{rs}(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}\mathrm{d}\mathbf{r}_1\mathrm{d}\mathbf{r}_2. \tag{3.25}$$

The γ_{HF} coefficient in Eqs. (3.21) and (3.24) is often employed in DFT approximations, where $\gamma_{\mathrm{HF}} \neq 0$ are denoted hybrid functionals since they include a fraction of HF-type exchange (obtained with KS orbitals), mimicking the Hartree–Fock scheme. These models will also modify the extent of the exchange obtained from the exchange–correlation potential (see Table 3.2). We discuss approximate exchange–correlation functionals further in Section 3.3.1 below but γ_{HF} is included here for completion.

3.3.1 Density Functional Approximations

The KS theory is exact if no approximations are introduced in $E_{xc}[\rho]$, even when employing a single-determinant reference state. We can formally write

$$E_{xc}[\rho] = \int e_{xc}(\rho)\, d\mathbf{r}, \tag{3.26}$$

where $e_{xc}(\rho)$ is the exchange-correlation energy density. The exact form of $E_{xc}[\rho]$, and hence of $\epsilon_{xc}(\rho)$ would ensure the equality in Eq. (3.13), and provide the exact energy of any given N-electron system. Unfortunately, the exact form of this functional is generally unknown, and approximations must be employed in practice. The first entries of Table 3.2 show commonly employed approximations (the lower entries are functionals modified for multi-configurational wave functions discussed in the next sections). The simplest functional forms are based on the local density approximation (LDA), constructed for a model system based on a uniform electron gas. In this case, the exact exchange functional is known, while there exist very accurate parameterizations for the correlation functional [57, 58]. The next steps are the generalized gradient approximation (GGA) and *meta*-GGA which additionally involves the gradient, ξ, and parts of the kinetic energy, τ

$$\xi(\mathbf{r}) = \nabla\rho(\mathbf{r}) \cdot \nabla\rho(\mathbf{r}) \tag{3.27}$$

$$\tau(\mathbf{r}) = \sum_{pq} (\nabla\phi_p(\mathbf{r}))(\nabla\phi_q(\mathbf{r}))^* D_{pq}. \tag{3.28}$$

The approximate exchange-correlation functional can now be written

$$E_{xc}[\rho, \xi, \tau] = \int e_{xc}(\rho, \xi, \tau) d\mathbf{r}. \tag{3.29}$$

Table 3.2 Common DFT exchange–correlation functionals for regular KS-DFT and for multi-configurational KS-DFT. The examples are composite functionals, i.e., they contain both exchange and correlation parts (an expression for E_x^{HF} in the hybrid functional is given in Eq. 3.21). For a shorter notation, the functionals have not been generalized to depend on spin.

Method	Input parameters	Examples
	Conventional	
LDA	$E_{xc}[\rho]$	VWN [57], PW92 [58]
GGA	$E_{xc}[\rho, \xi]$	BLYP [60, 61], PW91 [62], PBE [63]
Hybrid	$(1-\gamma_{HF})E_x[\rho,\xi] + \gamma_{HF}E_x^{HF}[\{\phi_i\}] + E_c[\rho,\xi]$	PBE0, B3LYP [64]
meta-GGA	$E_{xc}[\rho, \xi, \tau]$	TPSS [65]
	On-top density (multi-configurational)	
LDA	$E_{xc}^{sr}[\rho; P_2]$	tVWN [34],
GGA	$E_{xc}^{ot}[\rho, \xi, P_2]$, $E_{xc}^{ot}[\rho, \xi, P_2, \nabla P_2]$	tPBE [34], tBLYP [34], ftPBE [66], ftBLYP [66]
	Range-separated (multi-configurational)	
LDA	$E_{xc}^{sr}[\rho; \mu]$	sr-VWN [36, 49, 67, 68], sr-PW92 [69]
GGA	$E_{xc}^{sr}[\rho, \xi; \mu]$	PBEHSE(RI) [51,70–72], PBEGWS [69]
Hybrid	$(1-\gamma_{HF})E_x^{sr}[\rho,\xi;\mu] + \gamma_{HF}E_x^{sr\text{-}HF}[\{\phi_i\};\mu] + E_c[\rho,\xi;\mu]$	PBE0-GWS
meta-GGA	$E_{xc}[\rho, \xi, \tau; \mu]$	TPSS-GEMH [73]

So far, we have not included any reference to electron spin in the exchange–correlation functionals; introducing such a dependence is not necessary from a theoretical point of view where the exact energy can be solely defined from the total density. Yet, it has turned out that approximate semi-local functionals must depend on the spin density (and its gradients) to be accurate for open-shell systems [59]. We thus define a general approximate functional as

$$E_{\mathrm{xc}}[\rho^{\alpha},\rho^{\beta},\ldots] = \int e_{\mathrm{xc}}(\rho^{\alpha},\xi^{\alpha},\xi^{\beta},\xi^{\alpha\beta}\ \cdots)\mathrm{d}\mathbf{r}, \tag{3.30}$$

where $\xi^{\sigma\sigma'}$ is defined

$$\xi^{\sigma\sigma'} = \nabla\rho^{\sigma}(\mathbf{r})\cdot\nabla\rho^{\sigma'}(\mathbf{r}). \tag{3.31}$$

The gradient or spin-gradient in Eqs. (3.27) and (3.31) are calculated as

$$\nabla\rho^{\sigma} = \sum_{pq}\bar{D}_{p\sigma q\sigma}\nabla\Omega_{pq}. \tag{3.32}$$

The τ in Eq. (3.28) can be generalized to be spin-dependent in a similar manner. For open-shell systems, an alternative to Eqs. (3.31)–(3.32) is to employ the spin-density ρ^{s} and its gradients (cf. Eq. 3.16).

3.3.2 Density Functional Theory for Excited States

The by far most common method to calculate excited states and excitation energies with DFT is the TD-DFT formalism [1, 74, 75]. In practical implementations, the TD-DFT formulations employ a linear response framework, originally formulated for wave function theory [76–78]. Since TD-DFT is treated in more detail in Chapter 2, we will not provide a detailed derivation of the underlying equations; here, we just note that the TD-DFT equations in second-quantization can be formulated by means of time-dependent orbital rotation operators

$$\hat{\kappa}(t) = \sum_{pq}(\kappa_{pq}(t)\hat{E}_{pq} + \kappa^{*}_{pq}(t)\hat{E}_{qp}). \tag{3.33}$$

We only investigate excitations resulting in singlet states here, but the formalism can be extended to higher multiplicities, which requires more general orbital rotation operators. When applied as in Eq. (3.14), $\hat{\kappa}(t)$ can be employed to describe the time-development of $|0\rangle$ ultimately leading to the eigenvalue problem [21] given in Chapter 2. Here we provide a more general form of this equation

$$\left[\begin{pmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{B}^{*} & \mathbf{A}^{*}\end{pmatrix} - \omega\begin{pmatrix}\boldsymbol{\Sigma} & \boldsymbol{\Delta}\\ -\boldsymbol{\Delta}^{*} & -\boldsymbol{\Sigma}^{*}\end{pmatrix}\right]\begin{pmatrix}\mathbf{X}\\ \mathbf{Y}^{*}\end{pmatrix} = \mathbf{0} \tag{3.34}$$

which can also accommodate a multi-configurational wave function, as required in the subsequent sections. Staying now in a single reference framework, the $\mathbf{A}$ and $\mathbf{B}$ matrices become in our second-quantization formulation

$$A_{pq,rs} = \langle 0|[\hat{E}_{pq},[\hat{f}^{\mathrm{KS}}_{0},\hat{E}^{\dagger}_{rs}]]|0\rangle + \langle 0|[\hat{E}_{pq},(\hat{f}^{\omega}_{\mathrm{Hxc},rs})^{\dagger}]|0\rangle \tag{3.35}$$

$$B_{pq,rs} = \langle 0|[\hat{E}_{pq},[\hat{f}^{\mathrm{KS}}_{0},\hat{E}_{rs}]]|0\rangle + \langle 0|[\hat{E}_{pq},\hat{f}^{\omega}_{\mathrm{Hxc},rs}]|0\rangle, \tag{3.36}$$

where, $\hat{f}^{\mathrm{KS}}_{0}$ is the unperturbed Kohn–Sham operator, and $\hat{f}^{\omega}_{\mathrm{Hxc}}$ contains the Hartree-exchange-correlation kernel (cf. Chapter 2). The $\mathbf{X}$ and $\mathbf{Y}$ vectors contain the (Fourier transformed) $\boldsymbol{\kappa}$ coefficients from Eq. (3.33). Finally, the structures of the $\boldsymbol{\Sigma}$ and $\boldsymbol{\Delta}$ matrices are

$$\boldsymbol{\Sigma}_{pq,rs} = \langle 0|[\hat{E}_{pq},\hat{E}^{\dagger}_{rs}]|0\rangle \tag{3.37}$$

$$\mathbf{\Delta}_{pq,rs} = \langle 0|[\hat{E}_{pq}, \hat{E}_{rs}]|0\rangle. \tag{3.38}$$

For regular KS TD-DFT, $\mathbf{\Sigma}$ reduce to a simple diagonal matrix with 1 or -1 on the diagonal while $\mathbf{\Delta}$ vanish. Employing this simple diagonal form in Eq. (3.34) leads to the form given in Chapter 2, but note that a similar reduction does not occur for multi-configurational wave functions. For further details on the underlying equations for TD-DFT we refer to Chapter 2 as well as Refs. [55, 78, 79] concerning the second-quantization formulation of TD-DFT.

The equations in Eqs. (3.34)–(3.38) are the core equations employed in TD-DFT for calculation of excitation energies and oscillator strengths. From Eqs. (3.35)–(3.38) we infer that this set of equations are expressed with the ground-state orbitals as the basis, and that the orbital rotation operator is employed to generate this basis. This has some consequences for which processes can be described. Moreover, the approximations employed in the construction of the exchange-correlation functionals also have consequences, and we investigate these consquences below. In Section 3.4 on multi-configurational DFT, we will see that some of the restrictions from conventional (TD)-DFT is lifted by the use of a multi-configurational wave function.

3.3.2.1 Issues Within the Time-Dependent Density Functional Theory Ansatz

The orbital rotation operator used to generate a basis for the TD-DFT equations (see Eqs. (3.34, 3.38) above) only generates configurations that are singly excited with respect to the ground state. Therefore, higher order excitation classes such as excitations with a high degree of double excitation character are not included in TD-DFT, unless they are described through the exchange–correlation functional. This is not the case within the adiabatic approximation, [7, 8, 80, 81] which is the standard approach in most quantum chemistry codes today. Using ground-state orbitals to generate a basis also means that excited states with orbitals very different from the ground-state are likely not to be described well. Thus, TD-DFT lack orbital relaxation effects which are included in state-specific (and to some degree also in state-average) optimization procedures. An example where orbital relaxation has a large effect is core-spectroscopy. However, large effects have also occasionally been observed for valence excitations (see e.g., Ref. [82]). Moreover, the linear response formulations only obtain excitation energies (and possibly oscillator strengths), but the actual wave function for the excited state is not obtained. Hence, the formalism is not optimal to investigate properties that requires this wave function directly. Finally, for compounds with significant multi-configurational character, the TD-DFT formalism is likely to fail (as is the case for the ground-state).

3.3.2.2 Self-Interaction Error

Without modifications, the KS scheme is plagued by a spurious self-interaction error, which can be inferred by comparing the KS equations to the corresponding HF scheme. Both the KS and HF methods employ a single Slater determinant, and the schemes have an overall similar structure (cf. Eq. 3.22). In the HF scheme, the KS operator is replaced by the Fock operator

$$\hat{f} = \sum_{pq}\sum_{\sigma} f_{pq}\hat{a}^{\dagger}_{p\sigma}\hat{a}_{q\sigma}, \tag{3.39}$$

where the matrix elements are given

$$f_{pq} = h_{pq} + \sum_{rs}\sum_{\sigma}((pq|rs) - (pr|sq))\bar{D}^{\mathrm{SCF}}_{r\sigma s\sigma}. \tag{3.40}$$

As for Eq. (3.24), the expression for $\bar{\mathbf{D}}^{\mathrm{SCF}}$ has been kept so that we can directly compare the two equations. Comparing Eqs. (3.24) and (3.40) reveals an important difference between Hartree–Fock

theory: for a single-electron ($p = q = r = s$), the two-electron terms cancel exactly in Hartree–Fock theory. This is not the case in KS-DFT, since the exchange part of $\hat{v}_{xc}$ does not exactly match the part containing the two-electron integrals. Accordingly, an electron can interact with itself in DFT. This unphysical behavior can naturally cause issues both in ground and excited state calculations. For calculation of excited states, one consequence of the self-interaction error is that Rydberg and CT excitations are often severely underestimated by TD-DFT [9–13]. To amend this behavior, an often employed method is to replace the two-electron operator with a range-separated variant

$$\hat{g}_{ij} = \hat{g}_{ij}^{\mathrm{lr}} + \hat{g}_{ij}^{\mathrm{sr}}. \tag{3.41}$$

The long-range part is described exact (i.e., using KS orbitals with the HF expression), whereas the short-range part is described through an exchange functional. The function controlling the separation has often been taken to be the error function

$$\hat{g}_{ij}^{\mathrm{lr}} = \frac{\mathrm{erf}(\mu r_{ij})}{|\mathbf{r}_i - \mathbf{r}_j|} \quad \text{and} \quad \hat{g}_{ij}^{\mathrm{sr}} = \frac{1 - \mathrm{erf}(\mu r_{ij})}{|\mathbf{r}_i - \mathbf{r}_j|}, \tag{3.42}$$

where μ is the range-separation parameter, given in reciprocal Bohr. In this chapter, we focus exclusively on the error function to control the range-separation (as in Eq. 3.42), but it should be noted that the definition of $\hat{g}_{ij}^{\mathrm{lr}}$ and $\hat{g}_{ij}^{\mathrm{sr}}$ can differ [72]. The procedure in Eqs. (3.41) leads to so-called range-separated hybrid functionals [83] such as CAM-B3LYP [84]. The range-separation technique has also been employed to merge multi-configurational wave functions and DFT, and we return to this in Section 4.3, where we also discuss the range-separation parameter.

3.3.2.3 Degeneracies, Near-Degeneracies and the Symmetry Dilemma

Many of the problems in DFT and TD-DFT are closely associated with the definition of a single Slater determinant for the non-interacting reference system. Degeneracies imposed by spatial or spin symmetry are often not possible to represent correctly with this wave function *ansatz*; an example is the introduction of spin in DFT where it is only possible to describe the spin component with $M_S = S$ or $M_S = -S$ within a given spin multiplet. However, the electronic state of interest may have an intermediate M_S value $-S < M_S < S$. Such a state can only be represented by a determinant of broken spin-symmetry, where the spin-densities become unphysical (see e.g., Ref. [85, 86]). The best known example is probably the H_2 dissociation curve where a broken-symmetry solution has lower energy for elongated H−H distances. Peculiarly, the dissociation curves of H_2 and several other small molecules with spin-dependent LDA and GGA functionals from a broken spin-symmetry solution give energies in good correspondence with experiment [59, 87] while large errors are obtained when the correct spin-symmetry is imposed. This has been denoted as the symmetry-dilemma of DFT [88] and the same dilemma exists for the Hartree–Fock theory [89, 90]. The dilemma can also occur for equilibrium structures as, e.g., in Cr_2 [91, 92]. A slightly different aspect of the dilemma is that spin in DFT can be formulated from either open-shell spin-restricted or spin-unrestricted single-Slater determinants and this has consequences for the definition of the exchange–correlation functionals. Yet, none of the present approximations takes this into account (see Jacob and Reiher [93] for an in-depth discussion of this aspect).

While the symmetry-dilemma in DFT may seem to be a problem for ground-states, it naturally extends to excited states if DFT behaves unphysical in the ground-state. In such cases, TD-DFT is expected to be erratic. In the next section, we investigate methods that attempt to fix such wrong behavior by incorporation of a multi-configurational wave function into DFT (and TD-DFT).

3.4 Multi-Configurational Density Functional Theory

With the definitions in the previous section, we can now formulate a generic energy expression for multi-configurational DFT. We define a multi-configurational wave function, parameterized as

$$|\Psi^{\mathrm{mc}}\rangle = e^{-\hat{\kappa}} \left(\sum_j c_j |j\rangle \right), \tag{3.43}$$

where $\mathbf{c}$ contains the configuration interaction coefficients and the orbital rotation operator were defined in Eq. (3.15). The density as well as one- and two-particle density matrices evaluated from this wave function are denoted $\rho^{\mathrm{mc}}(\mathbf{r})$, $\mathbf{D}^{\mathrm{mc}}$ and $\mathbf{\Gamma}^{\mathrm{mc}}$ (and similarly for the corresponding spin densities and matrices). The matrices can be calculated according to Eqs. (3.5), (3.6) and (3.18). The energy can now be defined,

$$\begin{aligned} E[\rho^{\mathrm{mc}}_\alpha, \rho^{\mathrm{mc}}_\beta, \cdots] &= E'_{\mathrm{mc}} + E'_{\mathrm{H}}[\rho^{\mathrm{mc}}] + E_{\mathrm{xc}}[\rho^{\mathrm{mc}}_\alpha, \rho^{\mathrm{mc}}_\beta, \cdots] \\ &= E'_{\mathrm{mc}} + \frac{1}{2} \sum_{pqrs} (pq|rs)' D^{\mathrm{mc}}_{pq} D^{\mathrm{mc}}_{rs} + E_{\mathrm{xc}}[\rho^{\mathrm{mc}}_\alpha, \rho^{\mathrm{mc}}_\beta, \cdots]. \end{aligned} \tag{3.44}$$

The energy expressions for E'_{mc} is typically similar to the usual expression for a multi-configurational wave function and also $E'_{\mathrm{H}}[\rho^{\mathrm{mc}}]$ will generally have the same forms as in DFT (cf. Eq. 3.11). However, since $\mathbf{D}^{\mathrm{SCF}}$ is replaced by $\mathbf{D}^{\mathrm{mc}}$ the expressions are typically modified in some way to avoid double-counting of electron correlation. Therefore, the two first terms are denoted with a prime. The approximate functional is denoted $E_{\mathrm{xc}}[\rho^{\mathrm{mc}}_\alpha, \rho^{\mathrm{mc}}_\beta, \ldots]$ and will often (but not always) differ from the functional approximations in Eq. (3.30) which are made to comply with single-determinant reference states. The modifications made in the individual terms in Eq. (3.44) are different in different models, and not all models modify all terms. These details are discussed below with outset in three different multi-configurational DFT methods.

3.4.1 Semi-Empirical Multi-Configurational Density Functional Theory

The simplest hybrids between DFT and multi-configurational wave functions are semi-empirical. These methods typically employ conventional DFT functionals with a scaling factor to reduce the issue with double-counting of the electron correlation. A typical scheme do not calculate $E'_{\mathrm{H}}[\rho^{\mathrm{mc}}]$ and only employs the correlation part of E_{mc} from conventional approximations in Eq. (3.44). This correlation part is then scaled accordingly. Early work in this direction was done by Lie and Clementi [94, 95] who employed the natural occupation numbers to scale the correlation energy. Later work has defined different scaling schemes [96]. We focus here on one of these later models, namely the MRCI/DFT method defined by Grimme and Waletzke [32]. We follow the derivation in the original paper, and note that the HF energy corresponding to Eq. (3.40) can be written

$$E_{\mathrm{HF}} = \sum_{pq} h_{pq} D^{\mathrm{SCF}}_{pq} + \frac{1}{2} \sum_{pqrs} (pq|rs) \Gamma^{\mathrm{SCF}}_{pqrs}. \tag{3.45}$$

Following Ref. [32] we focus on a closed-shell system where the only non-zero elements of the one-particle density matrix are inactive orbitals. In general, the one-particle density matrix is then given $\bar{D}^{\mathrm{SCF}}_{i\sigma j\sigma} = \delta_{ij}$ and $\bar{D}_{i\sigma i\sigma} = \bar{\omega}_i$, where $\bar{\omega}_i$ is the occupation number. For a closed-shell system, $D^{\mathrm{SCF}}_{ij} = 2\delta_{ij}$ and $\omega_i = \bar{\omega}_{i\alpha} + \bar{\omega}_{i\beta}$ are always 2. Moreover, the two-electron density matrix can be reduced to $\Gamma^{\mathrm{SCF}}_{ijkl} = D^{\mathrm{SCF}}_{ij} D^{\mathrm{SCF}}_{kl} - \frac{1}{2} D^{\mathrm{SCF}}_{il} D^{\mathrm{SCF}}_{kj}$. The Fock-operator in Eq. (3.40) can now be simplified

$$f_{kl} = h_{kl} + \sum_i \left((kl|ii) - \frac{1}{2}(ki|il) \right) \omega_k, \tag{3.46}$$

and the energy in Eq. (3.45) can be expressed

$$E_{\mathrm{HF}} = \sum_i f_{ii}\omega_i - \frac{1}{2}\sum_{ik}\left((ii|kk) - \frac{1}{2}(ik|ki)\right)\omega_i\omega_k. \tag{3.47}$$

The HF energy can be related to a given configuration in a configurational interaction (CI) Hamiltonian by the effective Hamiltonian [97]

$$\begin{aligned}\hat{H} - E_{\mathrm{HF}} = &-\sum_i f_{ii}\omega_i - \frac{1}{2}\sum_{ik}\left((ii|kk) - \frac{1}{2}(ik|ki)\right)\omega_i\omega_k \\ &+ \sum_{ij} f_{ij}\hat{E}_{ij} - \sum_{ijk}\left((ij|kk) - \frac{1}{2}(ik|jk)\right)\omega_k\hat{E}_{ij} + \frac{1}{2}\sum_{ijkl}(ij|kl)\hat{e}_{ijkl},\end{aligned} \tag{3.48}$$

which can be applied to construct an effective MRCI/DFT Hamiltonian. Grimme and Waletzke then define the Hamiltonian for diagonal elements as [32]

$$\begin{aligned}\langle K|\hat{H}^{\mathrm{MRCI/DFT}} - E^{\mathrm{DFT}}|K\rangle &= \langle K|\hat{H} - E_{\mathrm{HF}}|K\rangle \\ &- \sum_i^{n_{exc}}(f_{ii} - f_{ii}^{KS}) + \sum_i^{n_{exc}}(f_{aa} - f_{aa}^{KS}) + \frac{1}{n_{exc}}\sum_{ai}^{n_{exc}}(p_J(aa|ii) - p[N_o](ai|ai)),\end{aligned} \tag{3.49}$$

where the first term involves the expression from Eq. (3.48), and the remaining terms are DFT corrections. Notably, the elements f_{aa}^{KS} are KS orbital energies, while p_J and $p[N_o]$ are fitting constants. As noted in Ref. [98], the nomenclature can be slightly misleading since f_{aa} and E_{HF} are not truly HF orbital and total energies, respectively. Rather they are calculated as HF-like expressions (Eqs. (3.46) Eq. (3.47)), but with KS orbitals. The n_{exc} denotes the excitation level in the parent configuration. Concerning the off-diagonal elements, these are scaled by an exponential factor

$$\langle K|\hat{H}^{\mathrm{MRCI/DFT}}|L\rangle = \langle K|\hat{H}|L\rangle p_1 \exp(-p_2(\Delta E_{KL})^4), \tag{3.50}$$

where p_1 and p_2 again are scaling factors and ΔE_{KL} is the energy difference between the two configurations. The MRCI/DFT method generally calculates the KS matrix elements using a full exchange-correlation functional (i.e., with both exchange and correlation parts). The original MRCI/DFT method was optimized to the BHLYP functional [32]. More recent developments for the MRCI/DFT method has been carried out by Beck et al. [99] and the group of Marian [98, 100], re-designing the MRCI/DFT Hamiltonian to employ less empirical corrections than the original version.

Excited states and excitation energies in MRCI/DFT are naturally included through optimization of several roots during the wave function optimization, either through state-specific or state-average procedures (cf. Table 3.1).

3.4.2 Multi-Configurational Density Functional Theory Based the On-Top Pair Density

The DFT models discussed so far employ the electron density as a basic variable. However, higher-order density matrices can also be employed. In fact, the concept of a N-electron density matrix in quantum chemistry is rather old [89, 101], and has also been used prior to its introduction in quantum chemistry, e.g., in statistical mechanics and thermodynamics [102]. It seems a natural next step to expand approximate DFT functionals to higher-order density matrices. In these functionals, mutual inter-relations between density matrices are exploited to replace $\rho^{\alpha}(\mathbf{r})$ and $\rho^{\beta}(\mathbf{r})$ with quantities that do not suffer from breaking of spin-symmetry. The relations applied here can be found in the general overviews of density matrices and their use in quantum chemistry

[53, 56, 101, 103]. Here we focus on a branch of methods which employ the so-called on-top pair density [88].

Below, we first define the on-top pair density. With focus on the MC-PDFT method defined by Truhlar and Gagliardi [34], we next discuss how the on-top pair density can be used in exchange-correlation functionals together with multi-configurational wave functions, and how this method treats electronically excited states.

3.4.2.1 Density Matrices and the On-Top Pair Density

In addition to the one-electron charge density in Eq. (3.10), we can define the two-electron density matrix, $\rho_2(\mathbf{r}_1, \mathbf{r}_2)$ as [104]

$$\rho_2(\mathbf{r}_1, \mathbf{r}_2) = N(N-1) \int \dots \int \Phi^*(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)\Phi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)\mathrm{d}\boldsymbol{\sigma}_1\mathrm{d}\boldsymbol{\sigma}_2\mathrm{d}\mathbf{x}_3 \cdots \mathrm{d}\mathbf{x}_N. \quad (3.51)$$

The densities can be expressed in terms of the diagonal elements of *reduced* density matrices $\gamma_1(\mathbf{x}_1; \mathbf{x}_1)$ and $\gamma_2(\mathbf{x}_1, \mathbf{x}_2; \mathbf{x}_1, \mathbf{x}_2)$

$$\rho(\mathbf{r}_1) = \int \gamma_1(\mathbf{x}_1, \mathbf{x}_1)\mathrm{d}\sigma_1 \quad (3.52)$$

$$\rho_2(\mathbf{r}_1, \mathbf{r}_2) = \iint \gamma_2(\mathbf{x}_1, \mathbf{x}_2)\mathrm{d}\sigma_1\sigma_2. \quad (3.53)$$

We define the reduced density matrices below in second quantization [53],

$$\gamma_1(\mathbf{x}_1, \mathbf{x}_1) = \sum_{pq}\sum_{\sigma\tau} \bar{D}_{p\sigma q\tau}\phi^*_{p\sigma}(\mathbf{x}_1)\phi_{q\tau}(\mathbf{x}_1) \quad (3.54)$$

$$\gamma_2(\mathbf{x}_1, \mathbf{x}_2; \mathbf{x}_1; \mathbf{x}_2) = \sum_{pqrs}\sum_{\sigma\tau\mu\nu} \bar{\Gamma}_{p\sigma q\tau r\mu s\nu}\phi^*_{p\sigma}(\mathbf{x}_1)\phi_{q\tau}(\mathbf{x}_1)\phi^*_{r\mu}(\mathbf{x}_2)\phi_{s\nu}(\mathbf{x}_2). \quad (3.55)$$

Thus, the spatial part in Eq. (3.53) is

$$\rho_2(\mathbf{r}_1, \mathbf{r}_2) = \langle 0|\hat{\rho}_2(\mathbf{r}_1, \mathbf{r}_2)|0\rangle = \sum_{pqrs} \Omega_{pq}(\mathbf{r}_1)\Omega_{rs}(\mathbf{r}_2)\Gamma_{pqrs}, \quad (3.56)$$

from which it can be inferred that $\hat{\rho}_2(\mathbf{r}_1, \mathbf{r}_2) = \sum_{pqrs}\Omega_{pq}(\mathbf{r}_2)\Omega_{rs}(\mathbf{r}_2)\hat{e}_{pqrs}$. Our main focus here is the on-top pair density, which is related to the two-electron reduced density matrix as

$$P_2(\mathbf{r}) = \iint_{\mathbf{r}_1=\mathbf{r}_2} \gamma_2(\mathbf{x}_1, \mathbf{x}_2; \mathbf{x}_1, \mathbf{x}_2)\mathrm{d}\boldsymbol{\sigma}_1\mathrm{d}\boldsymbol{\sigma}_2. \quad (3.57)$$

According to the definitions in Eqs. (3.56) and (3.57) we can write the on-top pair density as

$$P_2(\mathbf{r}) = \sum_{pqrs} \Omega_{pq}(\mathbf{r})\Omega_{rs}(\mathbf{r})\Gamma_{pqrs}. \quad (3.58)$$

Before we discuss the multi-configurational on-top pair density functional theory defined by Truhlar and Gagliardi [34], it should be noted that this method has a number of predecessors, building on similar ideas: many of these methods were proposed in the early to mid 2000's by Gräfenstein and Cremer [105, 106], Gusarov et al. [107] and Takeda et al. [108]. Even prior to these efforts, the use of on-top pair densities in DFT together with multi-configurational wave functions has been proposed. For instance, Moscarda and San-Fabian [109] already defined such a method in the beginning of the 1990s. Earlier still, Ziegler, Rauk and Baerends pointed out that atomic multiplet energies (which requires more than a single-determinant wave function) could not be obtained within regular KS–DFT [110]. In view of the discussion of Ziegler et al. [110], Perdew and co-workers discussed the pair-density in relation to strong correlation [88, 111],

while Becke et al. [112] discussed how to transform the spin-polarized LDA functional, depending on $\rho^{\alpha}(\mathbf{r})$ and $\rho^{\beta}(\mathbf{r})$ to functionals depending on $\rho(\mathbf{r})$ and $P_2(\mathbf{r})$.

It is instructive for the discussion of the MC-PDFT method of Ref. [34] to summarize how $\rho^{s}(\mathbf{r})$, $\rho^{\alpha}(\mathbf{r})$ and $\rho^{\beta}(\mathbf{r})$ are related to $P_2(\mathbf{r})$. As noted previously in this chapter, the one- and two-particle density matrices significantly simplify for a single-determinant wave function. In this case it can be shown that $\gamma_2(\mathbf{x}_1\mathbf{x}_2;\mathbf{x}_1\mathbf{x}_2)$ can be resolved as

$$\gamma_2(\mathbf{x}_1,\mathbf{x}_2;\mathbf{x}_1,\mathbf{x}_2) = \gamma_1(\mathbf{x}_1;\mathbf{x}_1)\gamma_1(\mathbf{x}_2;\mathbf{x}_2) - \gamma_1(\mathbf{x}_2;\mathbf{x}_1)\gamma_1(\mathbf{x}_1;\mathbf{x}_2). \tag{3.59}$$

Both $\gamma_1(\mathbf{x}_1;\mathbf{x}_2)$ and $\gamma_2(\mathbf{x}_1,\mathbf{x}_2;\mathbf{x}_1,\mathbf{x}_2)$ can further be resolved into α and β parts [56]. Here we only need the expression for $\gamma_1(\mathbf{x}_1;\mathbf{x}_2)$

$$\gamma_1(\mathbf{x}_1;\mathbf{x}_2) = \gamma_1^{\alpha\alpha}(\mathbf{x}_1;\mathbf{x}_1) + \gamma_1^{\beta\beta}(\mathbf{x}_1;\mathbf{x}_1). \tag{3.60}$$

By applying Eq. (3.60) when inserting Eq. (3.59) into the definition of $P_2(\mathbf{r})$ (cf. Eq. (3.57) or Eq. 3.58), the on-top pair density can be resolved (after carrying out the spin integration) into

$$P_2(\mathbf{r}) = 2\rho^{\alpha}(\mathbf{r})\rho^{\beta}(\mathbf{r}). \tag{3.61}$$

From Eq. (3.61) a relation between the spin-density $\rho^{s}(\mathbf{r})$ in Eq. (3.18) and the on-top pair density can be established

$$\rho^{s}(\mathbf{r}) = \rho(\mathbf{r})\left(1 - \frac{2P_2(\mathbf{r})}{\rho^2(\mathbf{r})}\right)^{\frac{1}{2}}. \tag{3.62}$$

The above expression is the one employed in Ref. [34] to define a multi-configurational DFT model. We continue discussing this model in the section below.

3.4.2.2 Energy Functional and Excited States with the On-Top Pair Density

With the definition of the on-top pair density in the previous section, we can now define a DFT functional with $\rho(\mathbf{r})$ and $P_2(\mathbf{r})$. The goal is to handle static correlation through a multi-configurational wave function, while the functional complies with the symmetry imposed by this wave function (and is not exposed to symmetry-breaking).

It should be noted, however, using present DFT approximations with an on-top pair density and a multi-configurational wave function in Eq. (3.44), will lead to double counting of electron correlation effects. This led Truhlar and Gagliardi [34] to suggest an MC-PDFT method where only the Coulomb energy and the one-electron terms of the Hamiltonian are obtained from the MCSCF wave function (reducing E'_{mc} in Eq. (3.44) to $E^{\mathrm{1e}}_{\mathrm{mc}}$, containing only the kinetic and electron-nuclear attraction parts). This avoids double counting and the remaining energy is obtained from an on-top pair density functional, for which the MCSCF wave function is employed to obtain $\rho(\mathbf{r})$ and $P_2(\mathbf{r})$, and possibly their gradients. Accordingly, the MC-PDFT energy is defined

$$\begin{aligned} E[\rho^{\mathrm{mc}}, P_2^{\mathrm{mc}}, \ldots] &= E^{\mathrm{1e}}_{\mathrm{mc}} + E_{\mathrm{H}}[\rho^{\mathrm{mc}}] + E^{\mathrm{ot}}_{\mathrm{xc}}[\rho^{\mathrm{mc}}, P_2^{\mathrm{mc}}, \ldots] \\ &= \sum_{pq} h_{pq} D^{\mathrm{mc}}_{pq} + \frac{1}{2}\sum_{pqrs}(pq|rs) D^{\mathrm{mc}}_{pq} D^{\mathrm{mc}}_{rs} + E^{\mathrm{ot}}_{\mathrm{xc}}[\rho^{\mathrm{mc}}, P_2^{\mathrm{mc}}, \ldots]. \end{aligned} \tag{3.63}$$

The superscript "mc" indicates (as in Eq. 3.44) that the density and on-top pair density are obtained through Eq. (3.18) and Eq. (3.58), i.e., the one- and two-particle density matrices are obtained from a multi-configurational wave function. Currently, a number of different multi-configurational wave functions have been combined with MC-PDFT, ranging from CASSCF [34] to RASSCF [113] and GASSCF [114]. Very recently, a DMRG variant was also reported [115].

While advantages over traditional DFT functionals are certainly seen for pair-density functionals, a few comments on their use in MC-PDFT are in order. First, it should be noted that E_{xc}^{ot} in Eq. (3.63) is evaluated as a correction *after* obtaining a density from a multi-configurational wave function; thus the energy is not obtained self-consistently. A similar procedure has previously been employed in models that translate the correlation part of E_{xc} in terms of $\rho(\mathbf{r})$ and $P_2(\mathbf{r})$ [116, 117]. Moreover, the manipulations used in the previous section to relate spin- and on-top pair densities (Eq. 3.62), assumed the wave function to be a single determinant. Hence, the use of Eq. (3.62) is not without issues for multi-configurational wave functions: values of $2P_2/\rho^2$ below 1 will generate complex values for $\rho^s(\mathbf{r})$. Further, although formal constraints on $P_2(\mathbf{r})$

$$0 \leqslant P_2(\mathbf{r}) \leqslant \frac{1}{2}\rho^2(\mathbf{r}), \tag{3.64}$$

can be derived, the upper limit $P_2(\mathbf{r}) = \frac{1}{2}\rho^2(\mathbf{r})$ is not valid for multi-configurational wave functions [112]. Hence, a procedure for these cases must be developed. One possibility is to extrapolate the functional to imaginary values; this route was taken by Mielich et al. [116]. Another method [116] is to redefine $P_2(\mathbf{r})$ to $P_2(\mathbf{r}) = 0$ or $P_2(\mathbf{r}) = \frac{\rho^2(\mathbf{r})}{2}$. The route taken in Ref. [34] uses the $P_2(\mathbf{r}) = 0$ strategy to translate existing functionals from conventional DFT directly according to

$$E_{xc}^{ot}[\rho^{mc}, P_2^{mc}] = E_{xc}^{ot}\left[\rho^{mc}, \begin{cases} \rho\left(1 - \dfrac{2P_2^{mc}}{\rho^2}\right)^{\frac{1}{2}} & \text{if } \left(\dfrac{2P_2^{mc}}{(\rho^{mc})^2}\right)^{\frac{1}{2}} \leq 1 \\ 0 & \text{if } \left(\dfrac{2P_2^{mc}}{(\rho^{mc})^2}\right)^{\frac{1}{2}} > 1 \end{cases}\right]. \tag{3.65}$$

We have for brevity only defined an LDA form, i.e., the functional only depends on $\rho^{mc}(\mathbf{r})$ and $P_2^{mc}(\mathbf{r})$. We here follow Ref. [34] and refer to Eq. (3.65) as an on-top pair density translated functional, denoted tDFT. For GGA functionals, one can either employ partly translated functionals, which only employ a gradient of $\rho(\mathbf{r})$ or fully translated functionals [66] with gradients of both $\rho(\mathbf{r})$ and $P_2(\mathbf{r})$. A number of examples are listed in Table 3.2.

The MC-PDFT method from Ref. [34] relies on either a state-specific or state-average method for excited states (cf. Table 3.1). Since the contribution from the exchange–correlation functional is evaluated after the wave function optimization, the method cannot easily be adapted to a linear response formalism in its current form. However, methods that include an on-top functional are not fundamentally restricted to state-specific (or state-average) optimization of excited states. If the underlying wave function is optimized along with the contribution from the wave function, they may be adapted to a linear response framework.

3.4.3 Multi-Configurational Density Functional Theory Based on Range-Separation

We previously saw that the range-separation of the two-electron operator (Eq. 3.41) can be applied to alter the amount of exact exchange, relieving the self-interaction error in approximate DFT functionals. The method of range-separation has also been suggested as a means to combine a multi-configurational wave function with DFT. In this way, the long-range interaction can be described by a multi-configurational wave function [36, 49]. The last method we discuss employs this strategy and is denoted multi-configurational short-range DFT (MC-srDFT) [50].

3.4.3.1 Energy Functional and Excited States in Range-Separated Methods

With a range-separated two-electron repulsion operator (cf. Eqs. 3.41 and 3.42), the resulting energy expression for a multi-configurational wave function is

$$\begin{aligned} E &= E_{\mathrm{mc}}^{\mathrm{lr}} + E_{\mathrm{H}}^{\mathrm{sr}}[\rho^{\mathrm{mc}}] + E_{\mathrm{xc}}^{\mathrm{sr}}[\rho_{\alpha}^{\mathrm{mc}}, \rho_{\beta}^{\mathrm{mc}}, \dots; \mu] \\ &= \langle \Psi^{\mathrm{mc}} | \hat{H}^{\mathrm{lr}} | \Psi^{\mathrm{mc}} \rangle + \frac{1}{2} \sum_{pqrs} (pq|rs)^{\mathrm{sr}} D_{pq}^{\mathrm{mc}} D_{rs}^{\mathrm{mc}} + E_{\mathrm{xc}}^{\mathrm{sr}}[\rho_{\alpha}^{\mathrm{mc}}, \rho_{\beta}^{\mathrm{mc}}, \dots; \mu]. \end{aligned} \tag{3.66}$$

The long-range Hamiltonian, $\hat{H}^{\mathrm{lr}}$, and the corresponding energy expression are identical to the regular MCSCF counterparts, but with the two-electron integrals replaced by integrals with the long-range part of $\hat{g}_{ij}$ in Eq. (3.42). Likewise, the Coulomb term, $E_{\mathrm{H}}^{\mathrm{sr}}[\rho^{\mathrm{mc}}]$, is evaluated with integrals over the short-range part of $\hat{g}_{ij}$ and hence denoted the short-range Coulomb interaction. The wave function $|\Psi^{\mathrm{mc}}\rangle$ is thus a multi-configurational wave function (cf. Eq. 3.43) for a fictitious system with long-range interaction $\hat{g}_{ij}^{\mathrm{lr}}$ and with the same ground-state density as the real system.

The range-separation procedure ensures that double counting of electron correlation is rigorously avoided. Unfortunately, it also means that approximate exchange-correlation functionals constructed for the regular two-electron repulsion operator cannot be straightforwardly employed. Tailored functionals are required, as they must be constructed explicitly for a modified two-electron interaction. We denote these functionals short-range DFT functionals (srDFT). Development of such functionals is still ongoing, but today both srLDA [36, 67], srGGA [69, 118] and *meta*-srGGA [73] variants have been developed (see Table 3.2).

Applications of MC–srDFT have so far mainly employed CASSCF type wave functions [52, 119–121] for which we use the acronym CAS–srDFT. However, both RASSCF, DMRG [122] and NEVPT2 [123] variants of multi-configurational wave functions have been implemented. These works extend previous works, combining srDFT functionals with CI and CC wave functions [69, 124]. In this context, it should be noted that although the wave function can be of any spin symmetry S, the current MC–srDFT model is restricted to wave functions corresponding to spin components $M_S = S$ or $M_S = -S$, since only these values give an appropriate spin density to input in the approximate exchange-correlation functional [37]. Therefore, investigations of dissociation energies have, as expected, led to considerable error [51] and investigation of excited states (or ground-states) with spin-couplings resulting in intermediate values of M_S are likewise expected to be problematic. In this regard, the MC-PDFT method is advantageous. It would, therefore, be interesting to combine the two methods. Such a method was in fact reported recently, although not with a multi-configurational wave function [125].

Concerning electronically excited states, the MC–srDFT method is unique among multi-configurational methods, as it is possible to obtain excited states through both a state-specific [50], state-average [126] and in a multi-configurational generalization of TD-DFT, which we here denote TD-MC–srDFT [127]. The latter gives rise to equations which are structurally equivalent to Eq. (3.34), but in which Eq. (3.34)–(3.38) are extended with configurational states and wave function parameters (for details, see Refs. [119, 127]). From these, it is possible to describe any rank of excitation with respect to the ground-state, unlike conventional TD-DFT. The TD-MC–srDFT variant of range-separated multi-configurational DFT is the model that has been most thoroughly investigated for excited states, and all results presented here (see Section 3.5) are calculated with this formalism.

3.4.3.2 The Range-Separation Parameter in Excited State Calculations

For multi-configurational DFT, the parameter μ defined together with Eq. (3.42) determines the separation between wave function theory and DFT. For the exact functional, the final

result will be μ-independent. However, for approximate functionals a dependence on μ enters. Therefore, the definition of an optimal μ parameter has been a goal for many studies in the field [51, 52, 128].

In an in-depth analysis for ground-state energetics of small molecules [51], it has been shown that an optimal μ parameter allocates as much dynamical correlation as possible within the DFT functional, while allowing the wave function to describe the remaining (static or dispersion) correlation effects. By comparison of HF–srDFT and MC–srDFT energies for both multi-configurational and single-reference systems, values in the vicinity of $\mu = 0.4$ was found to be optimal for this purpose. To investigate whether these conclusions could be transfered to excited states, a series of studies was recently carried out with MC–srDFT [52, 120, 122] and different values of μ for a set of organic molecules containing common classes of electronic transitions (e.g., $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ transitions). The set of molecules was chosen from the work by by Sauer, Thiel and co-workers [129–132] who have carried out extensive reference calculations with various electronic structure methods, including CASPT2, CC2, CC3 and DFT. The mean absolute deviations (MADs) for CAS–srDFT with respect to the other wave function methods are shown in Figure 3.1. The figure also shows the MAD against the compilation of best theoretical estimates from Ref. [129]. Note that in the limits of $\mu \rightarrow \infty$ and $\mu \rightarrow 0$, the MC–srDFT method reduces to pure MCSCF or pure DFT, respectively. These limits are also included in Figure 3.1.

From Figure 3.1, it is clear that values between $\mu = 0.3$ to $\mu = 0.5$ gives the best correspondence with the literature results. Hence, this range is also optimal for excitation energies (at least for the

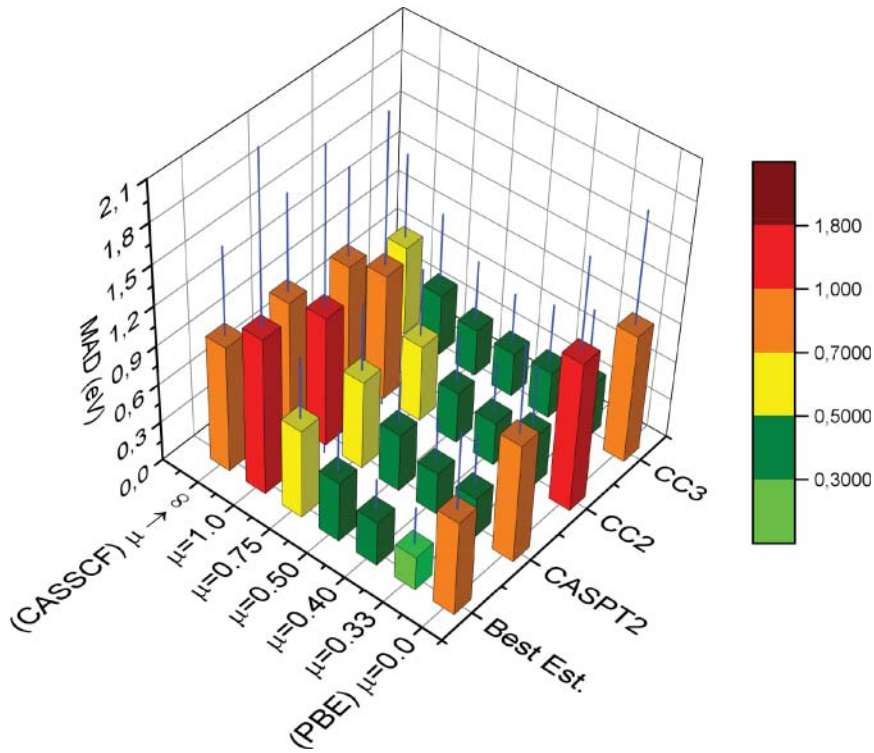


Figure 3.1 MADs for vertical excitation energies calculated with TD-CAS–srPBE with various μ-values against CC2, CC3, CASPT2 and the best estimates from Ref. [129]. Reproduced from Ref. [52].

set of organic molecules investigated). Initial studies for MnO_4^- suggest that similar values also work well for transition metal complexes, but here studies have been more sparse (see Section 3.5). The two methods MRCI/DFT and MC-PDFT have also been investigated for some of the molecules from Refs. [129–132]. We discuss these results in more detail below.

3.5 Illustrative Examples

In this section we focus on the performance of the three methods, MRCI/DFT, MC-PDFT and MC–srDFT. Our focus here is on performance for vertical excitation energies for which a wealth of data are available in the literature. We have selected a few recent studies on excited states of small to medium sized organic molecules. Although these systems in most cases do not posses multi-configurational ground states, they occasionally still can cause issues for TD-DFT. For instance, among the investigated molecules are linear polyenes whose 2^1A_g states are predominantly of double-excitation character. As seen from the discussion in the previous sections, this excitation class cannot be described by conventional TD-DFT.

In addition to simple organic molecules, we also compare the performance of the three methods for MnO_4^-, which is a notoriously difficult case, known to posses multi-configurational ground and excited states.

3.5.1 Excited States of Organic Molecules

In Table 3.3 we have compiled vertical excitation energies a few molecules selected from the larger investigation in Refs. [129–132]. Columns 3–6 contain results from CAS–srPBE [120, 122], CAS–tPBE [48], MRCI/DFT [130] and the conventional PBE functional [48]. First, it should be noted that the computational procedures are not entirely uniform for the methods in Table 3.3. Apart from occasional use of different basis sets, the CAS–srPBE excitation energies as well as excitation energies from PBE are calculated from a linear response formalism (i.e., TD-MC–srDFT and TD-DFT). The same is true for the reference CC2 excitation energies [129] in column 8, whereas the CASPT2 reference excitation energies (column 7) [129] are obtained from a state-average formalism. Similarly, the MRCI/DFT and CAS-tPBE results are obtained from state-average or state-specific optimizations. With this precaution in mind, we can comment on a few trends. Overall, all methods that combine DFT and a multi-configurational wave function compare better to the reference CC2 and CASPT2 values than pure DFT. Although the test set is still somewhat small, a trend that can be inferred is that PBE generally underestimates the excitation energies, compared to CASPT2 and CC2. The excitation energies generally increase when combining DFT with a multi-configurational wave function (for all three methods) and are in most cases closer to CASPT2 and CC2. In fairness, it should be mentioned that using a hybrid functional is also expected to result in higher excitation energies from TD-DFT.

A few of the systems deserve closer investigation. The $2^1A_g(\pi \rightarrow \pi^*)$ states in ethene, butadiene and hexatriene are examples of doubly excited states, which are not included in regular TD-DFT, neither for hybrid nor for non-hybrid functionals. Unfortunately, no CAS–tPBE results were reported for this state in Ref. [48]. The MRCI/DFT is slightly off the CASPT2 values, while CAS–srPBE is closer. For the 2^1A_g states also the reference values from CASPT2 and CC2 are rather different. The former is expected to be most accurate for excitations with a large degree of double excitation character.

Table 3.3 Excitation energies (in eV) from taken from the literature. The CAS–srDFT results are from Refs. [120, 122], the CAS–tPBE and PBE results from Ref. [48] and the MRCI/DFT results from Ref. [130]. The reference CASPT2 and CC2 values were taken from Ref. [129].

Molecule	State	CAS–srPBE	CAS–tPBE	MRCI/DFT	PBE	CASPT2	CC2
Ethene	$1^1B_{1u}(\pi \rightarrow \pi^*)$	7.87	6.77	7.96	7.35	8.62	8.40
Butadiene	$1^1B_u(\pi \rightarrow \pi^*)$	6.10	6.03	6.02	5.41	6.47	6.49
	$2^1A_g(\pi \rightarrow \pi^*)$	6.94		6.18		6.63	7.63
Hexatriene	$1^1B_u(\pi \rightarrow \pi^*)$	5.09	5.48	4.95	4.42	5.31	5.41
	$2^1A_g(\pi \rightarrow \pi^*)$	5.73		4.92		5.42	6.67
Benzene	$1^1B_{2u}(\pi \rightarrow \pi^*)$	5.87	5.09	5.04	5.14	5.05	5.27
Naphthalene	$1^1B_{3u}(\pi \rightarrow \pi^*)$	4.60	4.40	4.10	4.02	4.24	4.45
Furan	$1^1B_2(\pi \rightarrow \pi^*)$	6.42	6.51	6.33	5.87	6.43	6.75
Pyridine	$1^1B_1(n \rightarrow \pi^*)$	5.08	4.80	5.12	4.32	5.14	4.75
Pyrazine[a]	$1^1B_{3u}(n \rightarrow \pi^*)$	4.18	3.86	4.00	3.52	4.12	4.26
Pyrimidine	$1^1B_1(n \rightarrow \pi^*)$	4.61	4.06	4.36	3.75	4.44	4.49
Pyridazine	$1^1B_1(n \rightarrow \pi^*)$	3.83	3.24	3.63	3.11	3.78	3.90
s-Tetrazine	$1^1B_{3u}(n \rightarrow \pi^*)$	2.50	2.48	2.35	1.84	2.24	2.47
Formaldehyde	$1^1A_2(n \rightarrow \pi^*)$	3.81	3.55	3.71	3.77	3.99	4.09
Acetone	$1^1A_2(n \rightarrow \pi^*)$	4.31	4.35	4.23	4.20	4.44	4.52

The linear polyenes comprise an important class of organic molecules, and has been employed as simple model systems for carotenoids in photosynthetic pigments and light-sensor proteins such as rhodopsins. These biologically relevant systems also have low-lying excited states with double-excitation character [133]. Recent studies showed that MC–srDFT indeed could describe these states correctly [119]. Notably, part of these studies were carried out employing advanced embedding models to capture the effect of the protein [126].

Finally, we note that CT excitations have been investigated for both the MC-PDFT [113, 134] and MC–srDFT [119] methods, and both show marked improvement, compared to conventional DFT.

3.5.2 Excited States for a Transition Metal Complex

We now move to a more difficult system to illustrate the advantages of multi-configurational DFT. For this purpose, we chose the permanganate ion which is among the few transition metal complexes that has been investigated with all of the three multi-configurational DFT methods we consider here. The permanganate ion is a notorious multi-configurational system [135] as has been revealed by a number of theoretical studies on the absorption spectrum: a selection of studies is provided in Refs. [136–148] (see Ref. [144] for a short historical overview). The absorption spectrum has been the subject of some controversy: a well-resolved crystal spectrum was obtained in 1967 [149], but a number of different assignments [138–140, 142, 147] have been obtained from methods based on a single-reference wave function. Multi-configurational methods have been hampered by fairly large required active space, estimated to be 24 electrons in 17 orbitals [5, 150, 151]. A study on the MnO_4^- electronic absorption spectrum from 2013 managed to reach the required active space with RASPT2 [152] and we use this as reference here.

Table 3.4 Excitation energies (in eV) the lowest intense transitions of the MnO_4^- ion in vacuum from selected works. The CAS–srPBE, RAS–tPBE and MRCI/DFT energies are taken from Ref. [121], Ref. [155] and [154], respetively. The DFT (SOAP) and RASPT2 results are from Ref. [152]. The experimental values are from the crystal spectrum obtained in Ref. [149].

State	CAS(14,12)–srPBE	RAS(24,17)-tPBE	MRCI(10,10)/DFT	SOAP	RAS(24,17)PT2	Exp.
1^1T_2	2.21	2.22	2.80	3.06	2.33	2.2/2.3
2^1T_2	3.61	3.29	4.23	4.11	3.53	3.6
3^1T_2	4.01	3.90	4.84	4.97	4.20	4.1
4^1T_2	5.13	–	–	6.06	5.71	≈5.4

Several studies have also applied TD-DFT, but the method is noted to significantly overestimate the absolute excitation energies [141, 143, 144], although recent development in so-called constricted variational DFT seems to offer some improvement [148].

We have compiled results from the three multi-configurational DFT methods together with the RASPT2 and DFT results from Ref. [152] in Table 3.4. The experimental values are taken from the crystal spectrum [149]. As noted in Ref. [121] the marked vibronic progressions in the experimental UV-vis spectrum make the band-maxima somewhat ambiguous, which introduces some uncertainty in the comparison between theory and experiment. Slightly different experimental values therefore exist in the literature, e.g., the first state in Refs. [145, 148, 152] seems to correspond to the $\upsilon_0 \rightarrow \upsilon_2'$ peak, whereas others [137, 139, 140, 143, 144, 147, 153] report the band-maximum as the $\upsilon_0 \rightarrow \upsilon_1'$ peak. Following Ref. [121], we have reported both values in Table 3.4 although using any of the two values will not affect the conclusions: both CAS–srPBE and CAS–tPBE compare well to RASPT2 and to the experimental values. The two methods do not suffer from the overestimation seen for other DFT models. The MRCI/DFT results are in this particular case not a significant improvement to DFT, as they also overestimate the transitions somewhat. In Ref. [154], this was speculated to be due to a too small active space. In this regard, CAS-srDFT usually offers an advantage: unlike CAS-tPBE and CASPT2 (and the corresponding RAS variants) CAS–srDFT obtains dynamical and static correlation simultaneously, whereas both the CAS-tPBE and CASPT2 methods obtain dynamical correlation after initially optimization the CAS wave function. This difference means that CAS-srDFT often can be employed with much smaller CAS expansions, which is also evident from Table 3.4.

3.6 Outlook

This chapter has surveyed the current status in multi-configurational DFT with emphasis on three methods. A number of issues with conventional (single determinant) DFT for calculation of excited states have been summarized. Many of these issues may be solved by moving to a multi-configurational reference state. The main challenge when combining a multi-configurational wave function and DFT is avoiding double counting of electron correlation, while maintaining an efficient and accurate scheme. The three methods summarized here are chosen as they illustrate different strategies to mimimize the double counting issue. The first method (MRCI/DFT) is semi-empirical and uses a fitting procedure when combining an MRCI Hamiltonian with DFT. Although this does not rigorously remove double counting, it has turned out to work well for a number of cases, including cases where DFT usually fails. The two remaining schemes rely on

pair-density functionals or range-separation. While these schemes are more elaborate and do not rely on fitting, they also face larger challenges in their development. A particular challenge is that they require tailored functionals, and improving the employed functionals will still be an important branch of development over the coming years. Still, even with current functionals the methods have shown significant improvements over conventional DFT, both for excited states of organic chromophores and for transition metal complexes. Another important branch of development will be to extend the methods to the condensed phase. This work has been initiated for some of the methods, but is still in its infancy.

Acknowledgments

EDH thanks the Carlsberg foundation, the Villum Young Investigator Program (Grant No. 29412), and the European Commission (MetEmbed, project 745967) for funding.

References

1 Runge, E. and Gross, E.K.U. (1984). Density-functional theory for time-dependent systems. *Phys. Rev. Lett.* 52: 997–1000.
2 Marques, M.A.L. and Gross, E.K.U. (2004). Time-dependent density functional theory. *Annu. Rev. Phys. Chem.* 55: 427–455.
3 Casida, M.E. and Huix-Rotllant, M. (2012). Progress in time-dependent density-functional theory. *Annu. Rev. Phys. Chem.* 63: 287–323.
4 Adamo, C. and Jacquemin, D. (2013). The calculations of excited-state properties with time-dependent density functional theory. *Chem. Soc. Rev.* 42: 845–856.
5 Veryazov, V., Malmqvist, P.-Å., and Roos, B.O. (2011). How to select active space for multi-configurational quantum chemistry? *Int. J. Quantum Chem.* 111: 3329–3338.
6 Szalay, P.G., Müller, T., Gidofalvi, G. et al. (2012). Multiconfiguration self-consistent field and multireference configuration interaction methods and applications. *Chem. Rev.* 112: 108–181.
7 Maitra, N.T., Zhang, F., Cave, R.J., and Burke, K. (2004). Double excitations within time-dependent density functional theory linear response. *J. Chem. Phys.* 120: 5932–5937.
8 Elliot, P., Goldson, S., Canahui, C., and Maitra, N.T. (2011). Perspectives on double-excitations in TDDFT. *Chem. Phys.* 391: 110–119.
9 Tozer, D. and Handy, N.C. (1998). Improving virtual Kohn-Sham orbitals and eigenvalues: application to excitation energies and static polarizabilities. *J. Chem. Phys.* 109: 10180–10189.
10 Liao, M.-S., Lu, Y., and Scheiner, S. (2003). Performance assessment of density-functional methods for study of charge-transfer complexes. *J. Comput. Chem.* 24: 623–631.
11 Fabiano, E., Sala, F.D., Barbarella, G. et al. (2006). Optical properties of N-succinimidyl bithiophene and the effects of the binding to biomolecules: comparison between coupled-cluster and time-dependent density functional theory calculations and experiments. *J. Phys. Chem. B* 110: 18651–18660.
12 Perpetè, E., Preat, J., André, J.-M., and Jacquemin, D. (2006). An ab initio study of the absorption spectra of indirubin, isoindigo, and related derivatives. *J. Phys. Chem. A* 110: 5629–5635.
13 Dreuw, A. and Head-Gordon, M. (2004). Failure of time-dependent density functional theory for long-range charge transfer excited States: the zincbacteriochlorin-bacteriochlorin and bacteriochlorophyll-spheroidene complexes. *J. Am. Chem. Soc.* 126: 4007–4016.

14 Marti, K.H. and Reiher, M. (2010). The density matrix renormalization group algorithm in quantum chemistry. *Z. Phys. Chem.* 224: 583–599.

15 Chan, G.K.L. and Sharma, S. (2011). The density matrix renormalization group in quantum chemistry. *Annu. Rev. Phys. Chem.* 62: 465–481.

16 Kurashige, Y. (2014). Multireference electron correlation methods with density matrix renormalisation group reference functions. *Mol. Phys.* 112: 1485–1494.

17 Knecht, S., Hedegård, E.D., Keller, S. et al. (2016). New approaches for ab initio calculations of molecules with strong electron correlation. *Chimia* 70: 244–251.

18 Filippi, C., Assaraf, R., and Moroni, S. (2016). Simple formalism for efficient derivatives and multideterminant expansions in quantum Monte Carlo. *J. Chem. Phys.* 144: 194105.

19 Li Manni, G., Smart, S.D., and Alavi, A. (2016). Combining the complete active space self-consistent field method and the full configuration interaction quantum Monte Carlo within a super-CI framework, with application to challenging metal-porphyrins. *J. Chem. Theory Comput.* 12: 1245–1258.

20 Thomas, R.E., Sun, Q., Alavi, A., and Booth, G.H. (2015). Stochastic multiconfigurational self-consistent field theory. *J. Chem. Theory Comput.* 11: 5316–5325.

21 Olsen, J., Roos, B.O., Jørgensen, P., and Jensen, H.J.Å. (1988). Determinant based configuration-interaction algorithms for complete and restricted configuration-interaction spaces. *J. Chem. Phys.* 89: 2185–2192.

22 Fleig, T., Olsen, J., and Marian, C.M. (2001). The generalized active space concept for the relativistic treatment of electron correlation. I. Kramers-restricted two-component configuration interaction. *J. Comp. Phys.* 114: 4775.

23 Thyssen, J., Fleig, T., and Jensen, H.J.Å. (2008). A direct relativistic four-component multiconfiguration self-consistent-field method for molecules. *J. Comp. Phys.* 129: 034109.

24 Ma, D., LI Manni, G., and Gagliardi, L. (2011). The generalized active space concept in multiconfigurational self-consistent field methods. *J. Chem. Phys.* 135: 044128.

25 Andersson, K., Malmqvist, P.-Å., Roos, B.O. et al. (1990). Second-order perturbation theory with a CASSCF reference function. *J. Phys. Chem.* 94: 5483–5488.

26 Andersson, K., Malmqvist, P.-Å., and Roos, B.O. (1992). Second-order perturbation theory with a complete active space self-consistent field reference function. *J. Chem. Phys.* 96: 1218–1226.

27 Malmqvist, P.-Å., Pierloot, K., Shahi, A.R.M. et al. (2008). The restricted active space followed by second-order perturbation theory method: theory and application to the study of CuO_2 and Cu_2O_2 systems. *J. Chem. Phys.* 128: 204109.

28 Ma, D., Li Manni, G., Olsen, J., and Gagliardi, L. (2016). Second-order perturbation theory for generalized active space self-consistent-field wave functions. *J. Chem. Theory Comput.* 12: 3208–3213.

29 Angeli, C., Cimiraglia, R., Evangelisti, S. et al. (2001). Introduction of n-electron valence states for multireference perturbation theory. *J. Chem. Phys.* 114: 10252–10264.

30 Ghosh, S., Verma, P., Cramer, C.J. et al. (2018). Combining wave function methods with density functional theory for excited states. *Chem. Rev.* 118: 7249–7292.

31 Lischka, H., Nachtigallova, D., Aquino, A.J.A. et al. (2018). Multireference approaches for excited states of molecules. *Chem. Rev.* 118: 7293–7361.

32 Grimme, S. and Waletzke, M. (1999). A combination of Kohn-Sham density functional theory and multi-reference configuration-interaction methods. *J. Chem. Phys.* 111: 5645–5655.

33 Kleinschmidt, M., Marian, C.M., Waletzke, M., and Grimme, S. (2009). Parallel multireference configuration interaction calculations on mini-β-carotenes and β-carotene. *J. Chem. Phys.* 130: 044708.

34 Li Manni, G., Carlson, R.K., Luo, S. et al. (2014). Multiconfiguration pair-density functional theory. *J. Chem. Theory Comput.* 10: 3669–3680.

35 Gagliardi, L., Truhlar, D.G., Li Manni, G. et al. (2017). Multiconfiguration pair-density functional theory: a new way to treat strongly correlated systems. *Acc. Chem. Res.* 50: 66–73.

36 Savin, A. and Flad, H.-J. (1995). Density functionals for the Yukawa electron-electron interaction. *Int. J. Quantum Chem.* 56: 327–332.

37 Hedegård, E.D., Toulouse, J., and Jensen, H.J.Å. (2018). Multiconfigurational short-range density-functional theory for open-shell systems. *J. Chem. Phys.* 148: 214103.

38 Theophilou, A.K. (1972). The energy density functional formalism for excited states. *J. Phys. C* 12: 5419–5430.

39 Gross, E.K.U., Oliveira, L.N., and Kohn, W. (1988). Rayleigh-Ritz variational principle for ensembles of fractionally occupied states. *Phys. Rev. A* 37: 2805–2808.

40 Gross, E.K.U., Oliveira, L.N., and Kohn, W. (1988). Density-functional theory for ensembles of fractionally occupied states. I. Basic formalism. *Phys. Rev. A* 37: 2809–2820.

41 Pastorczak, E., Gidopoulos, N.I., and Pernal, K. (2013). Calculation of electronic excited states of molecules using the Helmholtz free-energy minimum principle. *Phys. Rev. A* 87: 062501.

42 Franck, O. and Fromager, E. (2014). Generalized adiabatic connection in ensemble density-functional theory for excited states: example of the H_2 molecule. *Mol. Phys.* 112: 1684–1701.

43 Filatov, M. and Shaik, S. (2014). A spin-restricted ensemble-referenced Kohn-Sham method and its application to diradicaloid situations. *Chem. Phys. Lett.* 304: 429–437.

44 Filatov, M., Martinez, T.J., and Kim, K.S. (2016). Using the GVB Ansatz to develop ensemble DFT method for describing multiple strongly correlated electron pairs. *Phys. Chem. Chem. Phys.* 18: 21040–21050.

45 Filatov, M., Liu, F., Kim, K.S., and Martínez, T.J. (2016). Self-consistent implementation of ensemble density functional theory method for multiple strongly correlated electron pairs. *J. Chem. Phys.* 145: 244104.

46 Filatov, M. (2013). Assessment of density functional methods for obtaining geometries at conical intersections in organic molecules. *J. Chem. Theory Comput.* 9: 4526–4541.

47 Marian, C.M. and Gilka, N. (2008). Performance of the density functional theory/multireference configuration interaction method on electronic excitation of extended-systems. *J. Chem. Theory Comput.* 4: 1501–1515.

48 Hoyer, C.E., Ghosh, S., Truhlar, D.G., and Gagliardi, L. (2016). Multiconfiguration pair-density functional theory is as accurate as CASPT2 for electronic excitation. *J. Phys. Chem. Lett.* 7: 586–591.

49 Savin, A. (1996). On degeneracy, near degeneracy and density functional theory. In: *Recent Developments of Modern Density Functional Theory* (ed. J.M. Seminario), 327–357. Amsterdam: Elsevier.

50 Pedersen, J.K. (2004). *Description of Correlation and Relativistic Effects in Calculations of Molecular Properties*, Ph.D. thesis, University of Southern Denmark.

51 Fromager, E., Toulouse, J., and Jensen, H.J.Å. (2007). On the universality of the long-/short range separation in multiconfigurational density-functional theory. *J. Chem. Phys.* 126: 074111.

52 Hubert, M., Jensen, H.J.Å., and Hedegård, E.D. (2016). Excitation spectra of nucleobases with multiconfigurational density functional theory. *J. Phys. Chem. A* 120: 36–43.

53 Helgaker, T., Jørgensen, P., and Olsen, J. (2004). *Molecular Electronic-Structure Theory*. Wiley.

54 Saue, T. and Helgaker, T. (2002). Four-component relativistic Kohn-Sham theory. *J. Comput. Chem.* 23: 814–823.

55 Sałek, P., Vahtras, O., Helgaker, T., and Ågren, H. (2002). Density-functional theory of linear and nonlinear time-dependent molecular properties. *J. Chem. Phys.* 117: 9630–9645.

56 Parr, R. and Yang, W. (1989). *Density-Functional Theory of Atoms and Molecules*. Oxford University Press.

57 Vosko, S.H., Wilk, L., and Nusair, M. (1980). Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can. J. Phys.* 58: 1200–1211.

58 Perdew, J.P. and Wang, Y. (1992). Accurate and simple analytic representation of the electron-gas correlation energy. *Phys. Rev. B* 45: 13244–13249.

59 Gunnarson, O. and Lundqvist, B.I. (1976). Exchange and correlation in atoms, molecules, and solids by the spin-density-functional formalism. *Phys. Rev. B* 13: 4274–4298.

60 Becke, A.D. (1988). Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* 38: 3098–3100.

61 Lee, C., Yang, W., and Parr, R.G. (1988). Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* 37: 785–789.

62 Perdew, J.P. (1991). Unified theory of exchange and correlation beyond the local density approximation. In: *Electronic Structure of Solids '91* (eds. P. Ziesche and H. Eschrig), 11–20. Berlin: Akademie Verlag.

63 Perdew, J.P., Burke, K., and Ernzerhof, M. (1996). Generalized gradient approximation made simple. *Phys. Rev. Lett.* 77: 3865–3868.

64 Becke, A.D. (1993). Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* 98: 5648–5652.

65 Tao, J., Perdew, J.P., Staroverov, and Scuseria, G.E. (2003). Climbing the density functional ladder: nonempirical meta-generalized gradient approximation designed for molecules and solids. *Phys. Rev. Lett.* 91: 146401.

66 Carlson, R.K., Truhlar, D.G., and Gagliardi, L. (2015). Multiconfiguration pair-density functional theory: a fully translated gradient approximation and its performance for transition metal dimers and the spectroscopy of $Re_2Cl_8^{2-}$. *J. Chem. Theory Comput.* 11: 4077–4085.

67 Toulouse, J., Savin, A., and Flad, H.-J. (2004). Short-range exchange-correlation energy of a uniform electron gas with modified electron-electron interaction. *Int. J. Quantum Chem.* 100: 1047–1056.

68 Paziani, S., Moroni, S., Gori-Giorgi, P., and Bachelet, G.B. (2006). Local-spin-density functional for multideterminant density functional theory. *Phys. Rev. B* 73: 155111.

69 Goll, E., Werner, H.-J., and Stoll, H. (2005). A short-range gradient-corrected density functional in long-range coupled-cluster calculations for rare gas dimers. *Phys. Chem. Chem. Phys.* 7: 3917–3923.

70 Heyd, J., Scuseria, G.E., and Ernzerhof, M. (2003). Hybrid functionals based on a screened Coulomb potential. *J. Chem. Phys.* 118: 8207–8215.

71 Heyd, J. and Scuseria, G.E. (2004). Assessment and validation of a screened Coulomb hybrid density functional. *J. Chem. Phys.* 120: 7274–7280.

72 Toulouse, J., Colonna, F., and Savin, A. (2004). Long-range-short-range separation of the electron-electron interaction in density-functional theory. *Phys. Rev. A* 70: 062505.

73 Goll, E., Ernst, M., Moegle-Hofacker, F., and Stoll, H. (2009). Development and assessment of a short-range meta-GGA functional. *J. Chem. Phys.* 130: 234112.

74 Casida, M.E. (1996). Time-dependent density functional response theory for molecules. In: *Recent Advances in Density Functional Methods. Part I*, vol. 1 (ed. D.P. Chong), 155–192. World Scientific.

75 Bauernschmitt, R. and Ahlrichs, R. (1995). Treatment of electronic excitations within the adiabatic approximation of time dependent density functional theory. *Chem. Phys. Lett.* 256: 454–464.

76 Linderberg, J. and Öhrn, Y. (1973). *Propagators in Quantum Chemistry*. London: Academic Press.

77 Zubarev, D.L. (1974). *Nonequilibrium Statistical Thermodynamics*. New York: Consultant Bureau.

78 Olsen, J. and Jørgensen, P. (1985). Linear and non-linear response functions for an exact state and for an MCSCF state. *J. Chem. Phys.* 82: 3235–3264.

79 Sałek, P., Helgaker, T., and Saue, T. (2005). Linear response at the 4-component relativistic density-functional level: application to the frequency-dependent dipole polarizability of Hg, AuH and PtH_2. *Chem. Phys.* 311: 187–201.

80 Neugebauer, J., Baerends, E.J., and Nooijen, M. (2004). Vibronic coupling and double excitations in linear response time-dependent density functional calculations: dipole-allowed states of N_2. *J. Chem. Phys.* 121: 6155–6166.

81 Burke, K. (2012). Perspective on density functional theory. *J. Chem. Phys.* 136: 150901.

82 Eriksen, J.J., Sauer, S.P.A., Mikkelsen, K.V. et al. (2013). Failures of TDDFT in describing the lowest intramolecular charge-transfer excitation in para-nitroaniline. *Mol. Phys.* 111: 1235–1248.

83 Tawada, Y., Tsuneda, T., and Yanagisawa, S. (2004). A long-range-corrected time-dependent density functional theory. *J. Chem. Phys.* 120: 8425–8433.

84 Yanai, T., Tew, D.P., and Handy, N.C. (2004). A new hybrid exchange-correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.* 393: 51–57.

85 Koch, W. and Holthausen, M.C. (2008). *A Chemist's Guide to Density Functional Theory*, 2 ed. Wiley.

86 Neese, F. (2004). Definition of corresponding orbitals and the diradical character in broken symmetry DFT calculations on spin coupled systems. *J. Phys. Chem. Solids* 65: 781–785.

87 Seminario, J.M. (1994). A study of small systems containing H and O atoms using nonlocal functionals: comparisons with ab initio and experiment. *Int. J. Quantum Chem.* S28: 655–666.

88 Perdew, J.P., Savin, A., and Burke, K. (1995). Escaping the symmetry dilemma through a pair-density of spin-density functional theory. *Phys. Rev. A* 51: 4531–4541.

89 Löwdin, P.-O. (1955). Quantum theory of many-particle systems. I. Physical interpretations by means of density matrices, natural spin-orbitals, and convergence problems in the method of configurational interaction. *Phys. Rev.* 97: 1474–1489.

90 Löwdin, P.-O. (1955). Quantum theory of many-particle systems. II. Study of the ordinary Hartree-Fock approximation. *Phys. Rev.* 97: 1474–1489.

91 Delley, B., Freeman, A.J., and Ellis, D.E. Metal-metal bonding in Cr-Cr and Mo-Mo dimers: another success of local spin-density theory. *Phys. Rev. Lett.* 50: 488–491.

92 Dunlap, B.I. (1983). $X\alpha$, Cr_2, and the symmetry dilemma. *Phys. Rev. A* 27: 2217–2219.

93 Jacob, C.R. and Reiher, M. (2012). Spin in density-functional theory. *Int. J. Quantum Chem.* 112: 3661–3684.

94 Lie, G.C. and Clement, E. (1974). Study of the electronic structure of molecules. XXI. Correlation energy corrections as a functional of the Hartree–Fock density and its application to the hydrides of the second row atoms. *J. Chem. Phys.* 60: 1275–1287.

95 Lie, G.C. and Clement, E. (1974). Study of the electronic structure of molecules. XXII. Correlation energy corrections as a functional of the Hartree Fock type density and its application to the homonuclear diatomic molecules of the second row atoms. *J. Chem. Phys.* 60: 1288–1296.

96 Malcolm, N.O. and McDouall, J.J. (1998). A simple scaling for combining multiconfigurational wave functions with density functionals. *Chem. Phys. Lett.* 282: 121–127.

97 Wetmore, R.W. and Segal, G.A. (1975). Efficient generation of configuration interaction elements. *Chem. Phys. Lett.* 36: 478–483.

98 Marian, C.M., Heil, A., and Kleinschmidt, M. (2019). The DFT/MRCI method. *WIREs Comput. Mol. Sci.* 9: e1394.

99 Beck, E.V., Stahlberg, E.A., Burggraf, L.W., and Blaudeau, J.-P. (2008). A graphical unitary group approach-based hybrid density functional theory multireference configuration interaction method. *Chem. Phys.* 349: 158–169.

100 Lyskov, I., Kleinschmidt, M., and Marian, C.M. (2016). Redesign of the DFT/MRCI Hamiltonian. *J. Chem. Phys.* 144: 034104.

101 McWeeny, R. (1959). The density matrix in many-electron quantum mechanics. I. Generalized product functions. Factorization and physical interpretation of the density matrices. *Proc. R. Soc. Lond. A* 253: 242–259.

102 McWeeny, R. (1960). Some recent advances in density matrix theory. *Rev. Mod. Phys.* 32: 335–369.

103 McWeeny, R. (2004). *Spins in Chemistry*. Dover.

104 McWeeny, R. and Sutcliffe, B.T. (1969). *Methods of Molecular Quantum Mechanics*. Academic Press London.

105 Gräfenstein, J. and Cremer, D. (2000). The combination of density functional theory with multiconfiguration methods–CAS-DFT. *Chem. Phys. Lett.* 316: 569–577.

106 Gräfenstein, J. and Cremer, D. (2005). Development of a CAS-DFT method covering nondynamical and dynamical electron correlation in a balanced way. *Mol. Phys.* 103: 279–308.

107 Gusarov, S., Malmqvist, P.-Å., and Lindh, R. (2004). Using on-top pair density for construction of correlation functionals for multideterminant wave functions. *Mol. Phys.* 102: 2207–2216.

108 Takeda, R., Yamanaka, S., and Yamaguchi, K. (2004). Approximate on-top pair density into one-body functions for CAS-DFT. *Chem. Phys. Lett.* 96: 463–473.

109 Moscarda, F. and San-Fabian, E. (1991). Density-functional formalism and the two-body problem. *Phys. Rev. A* 44: 1549–1553.

110 Ziegler, T., Rauk, A., and Baerends, E.J. (1977). On the calculation of multiplet energies by the Hartree-Fock- Slater method. *Theoret. Chim. Acta* 271: 261–271.

111 Perdew, J.P., Ernzerhof, M., Burke, K., and Savin, A. (1997). On-top pair-density interpretation of spin density functional theory, with applications to magnetism. *Int. J. Quantum Chem.* 61: 197–205.

112 Becke, A.D., Savin, A., and Stoll, H. (1995). Extension of the local-spin-density exchange-correlation approximation to multiplet states. *Theor. Chim. Acta* 91: 147–156.

113 Presti, D., Truhlar, D.G., and Gagliardi, L. (2018). Intramolecular charge transfer and local excitation in organic fluorescent photoredox catalysts explained by RASCI-PDFT. *J. Phys. Chem. C* 122: 12061–12070.

114 Ghosh, S., Cramer, C.J., Truhlar, D.G., and Gagliardi, L. (2017). Generalized-active-space pair-density functional theory: an efficient method to study large, strongly correlated, conjugated systems. *Chem. Sci.* 8: 2741–2750.

115 Sharma, P., Bernales, V., Knecht, S., Truhlar, D.G., and Gagliardi, L. (2019). Density matrix renormalization group pair-density functional theory (DMRG-PDFT): singlet-triplet gaps in polyacenes and polyacetylenes. *Chem. Sci.* 10: 1716–1723.

116 Miehlich, B., Stoll, H., and Savin, A. (1997). A correlation-energy density functional for multideterminantal wave functions. *Mol. Phys.* 91: 527–536.

117 McDouall, J.J. (2003). Combining two-body density functionals with multiconfigurational wave functions: diatomic molecules. *Mol. Phys.* 101: 361–371.

118 Goll, E., Werner, H.-J., Stoll, H. et al. (2006). A short-range gradient-corrected spin density functional in combination with long-range coupled-cluster methods: application to alkali-metal rare-gas dimers. *Chem. Phys.* 329: 276–282.

119 Hedegård, E.D., Heiden, F., Knecht, S. et al. (2013). Assessment of charge-transfer excitations with time-dependent, range-separated density functional theory based on long-range MP2 and multiconfigurational self-consistent field wave functions. *J. Chem. Phys.* 139: 184308.

120 Hubert, M., Hedegård, E.D., and Jensen, H.J.Å. (2016). Investigation of multiconfigurational short-range density functional theory for electronic excitations in organic molecules. *J. Chem. Theory Comput.* 12: 2203–2213.

121 Olsen, J.M.H. and Hedegård, E.D. (2017). Modeling the absorption spectrum of the permanganate ion in vacuum and in aqueous solution. *Phys. Chem. Chem. Phys.* 19: 15870–15875.

122 Hedegård, E.D. (2016). Assessment of oscillator strengths with multiconfigurational short-range density functional theory for electronic excitations in organic molecules. *Mol. Phys.* 115: 26–38.

123 Fromager, E., Cimiraglia, R., and Jensen, H.J.Å. (2010). Merging multireference perturbation and density-functional theories by means of range separation: potential curves for Be_2, Mg_2, and Ca_2. *Phys. Rev. A* 81: 024502.

124 Leininger, T., Stoll, H., Werner, H.-J., and Savin, A. (1997). Combining long-range configuration interaction with short-range density functionals. *Chem. Phys. Lett.* 275: 151–160.

125 Garza, A.J., Bulik, I.W., Henderson, T.M., and Scuseria, G.E. (2015). Range separated hybrids of pair coupled cluster doubles and density functionals. *Phys. Chem. Chem. Phys.* 17: 22412–22422.

126 Hedegård, E.D., Olsen, J.M.H., Knecht, S. et al. (2015). Polarizable embedding with a multiconfiguration short-range density functional theory linear response method. *J. Chem. Phys.* 142: 114113.

127 Fromager, E., Knecht, S., and Jensen, H.J.Å. (2013). Multi-configuration time-dependent density-functional theory based on range separation. *J. Chem. Phys.* 138: 084101.

128 Fromager, E., Réal, F., Wåhlin, P. et al. (2009). On the universality of the long-/short-range separation in multiconfigurational density-functional theory. II. Investigating f^0 actinide species. *J. Chem. Phys.* 131: 054107.

129 Schreiber, M., Silva-Junior, M.R., Sauer, S.P.A., and Thiel, W. (2008). Benchmarks for electronically excited states: CASPT2, CC2, CCSD, and CC3. *J. Chem. Phys.* 128: 134110.

130 Silva-Junior, M.R., Schreiber, M., Sauer, S.P.A., and Thiel, W. (2008). Benchmarks for electronically excited states: time-dependent density functional theory and density functional theory based multireference configuration interaction. *J. Chem. Phys.* 129: 104103.

131 Silva-Junior, M.R., Schreiber, M., Sauer, S.P.A., and Thiel, W. (2010). Basis set effects on coupled cluster benchmarks of electronically excited states: CC3, CCSDR(3) and CC2. *Mol. Phys.* 108: 453–465.

132 Silva-Junior, M.R., Schreiber, M., Sauer, S.P.A., and Thiel, W. (2010). Benchmarks of electronically excited states: basis set effects on CASPT2 results. *J. Chem. Phys.* 133: 174318.

133 König, C. and Neugebauer, J. (2012). Quantum chemical description of absorption properties and excited-state processes in photosynthetic systems. *ChemPhysChem* 13: 386–425.

134 Ghosh, S., Sonnenberger, A.L., Hoyer, C.E. et al. (2015). Multiconfiguration pair-density functional theory outperforms Kohn-Sham density functional theory and multireference perturbation theory for ground-state and excited-State charge transfer. *J. Chem. Theory Comput.* 11: 3643–3649.

135 Buijse, M.A. and Baerends, E.J. (1990). Analysis of nondynamical correlation in the metal-ligand bond. Pauli repulsion and orbital localization in MnO_4^-. *J. Chem. Phys.* 93: 4129–4141.

136 Wolfsberg, M. and Helmholz, L. (1952). The spectra and electronic structure of the tetrahedral ions MnO_4^-, CrO_4^-, and ClO_4^-. *J. Chem. Phys.* 20: 837–843.

137 Johnson, K. and Smith, F.J. (1971). Scattered-wave model for the electronic structure and optical properties of the permanganate ion. *Chem. Phys. Lett.* 10: 219–223.

138 Johansen, H. and Rettrup, S. (1983). Limited configuration interaction calculation of the optical spectrum for the permanganate ion. *Chem. Phys.* 74: 77–81.

139 Nakai, H., Ohmori, Y., and Nakatsuji, H. (1991). Theoretical study on the ground and excited states of MnO_4^-. *J. Chem. Phys.* 95: 8287–8291.

140 Nooijen, M. (1999). Combining coupled cluster and perturbation theory. *J. Chem. Phys.* 111: 10815–10826.

141 van Gisbergen, S.J.A., Groeneveld, J.A., Rosa, A. et al. (1999). Excitation energies for transition metal compounds from time-dependent density functional theory. Applications to MnO_4^-, $Ni(CO)_4$, and $Mn_2(CO)_{10}$. *J. Phys. Chem. A* 103: 6835–6844.

142 Nooijen, M. and Lotrich, V. (2000). Extended similarity transformed equation-of-motion coupled cluster theory (extended-STEOM-CC): applications to doubly excited states and transition metal compounds. *J. Chem. Phys.* 113: 494–507.

143 Boulet, P., Chermette, H., Daul, C. et al. (2001). Absorption spectra of several metal complexes revisited by the time-dependent density-functional theory-response theory formalism. *J. Phys. Chem. A* 105: 885–894.

144 Neugebauer, J., Baerends, E.J., and Nooijen, M. (2005). Vibronic structure of the permanganate absorption spectrum from time-dependent density functional calculations. *J. Phys. Chem. A* 109: 1168–1179.

145 Jose, L., Seth, M., and Ziegler, T. (2012). Molecular and vibrational structure of tetroxo d^0 metal complexes in their excited states. A study based on time-dependent density functional calculations and Franck-Condon theory. *J. Phys. Chem. A* 116: 1864–1876.

146 Ziegler, T. (2012). A chronicle about the development of electronic structure theories for transition metal complexes. *Struct. Bond.* 143: 1–38.

147 Almeida, N.M., McKinlay, R.G., and Paterson, M.J. (2015). Excited electronic states of MnO_4^-: challenges for wave function and density functional response theories. *Chem. Phys.* 446: 86–91.

148 Seidu, I., Krykunov, M., and Ziegler, T. (2015). Applications of time-dependent and time-independent density functional theory to electronic transitions in tetrahedral d^0 metal oxides. *J. Chem. Theory Comput.* 11: 4041–4053.

149 Holt, S.L. and Ballhausen, C. (1967). Low temperature absorption spectra of $KMnO_4$ in $KClO_4$. *Theor. Chim. Acta* 7: 313–320.

150 Hedegård, E.D., Jensen, H.J.Å., and Kongsted, J. (2014). Polarizable embedding based on multiconfigurational methods: current developments and the road ahead. *Int. J. Quantum Chem.* 114: 1102–1107.

151 Stein, C.J. and Reiher, M. (2016). Automated selection of active orbital spaces. *J. Chem. Theory Comput.* 12: 1760–1771.

152 Su, J., Xu, W.-H., Xu, C.-F. et al. (2013). Theoretical studies on the photoelectron and absorption spectra of MnO_4^- and TcO_4^-. *Inorg. Chem.* 52: 9867–9874.

153 Hillier, I.H. and Saunders, V.R. (1970). Ab initio molecular orbital calculations of the ground and excited state of the permanganate and chromate ions. *Proc. R. Soc. A* 320: 161–173.

154 Escudero, D. and Thiel, W. (2014). Theoretical studies on the photoelectron and absorption spectra of MnO_4^- and TcO_4^-. *J. Chem. Phys.* 140: 194105.

155 Sharma, P., Truhlar, D.G., and Gagliardi, L. (2018). Multiconfiguration pair-density functional theory investigation of the electronic spectrum of MnO_4^-. *J. Chem. Phys.* 148: 124305.

4

Equation-of-Motion Coupled-Cluster Models

Monika Musiał

Institute of Chemistry, University of Silesia in Katowice, Szkolna 9, 40-006 Katowice, Poland

Abstract

Various computational schemes hidden under the name EOM-CC (equation-of-motion coupled-cluster) create a path through which the CC theory is applied to investigate excited, ionized and electron-attached molecular systems. In this chapter we consider five realizations of the EOM-CC approach focused on the studies of electronic excitations (EE), ionization potentials (IP), electron-attached (EA) states, double ionization potentials (DIP) and double-electron-attached (DEA) states. The direct application of the considered methods allows to study electronic states of the reference system as well as those differing from the reference by one (IP, EA) or two (DIP, DEA) electrons. In addition we propose an indirect application of the EOM-CC schemes which in some cases may be more interesting than the direct one. Namely, when the open-shell system A is studied then we may adopt as the reference one of its charged analogues: A^+, A^-, A^{2+}, A^{2-}on condition that it represents the closed-shell structure implying using the restricted Hartree-Fock (HF) reference. Then to recover the data relevant to the neutral system A we need to apply the EA, IP, DEA or DIP variant of the EOM-CC scheme, respectively. This can be generalized in the following way: to study with the EOM-CC approach the system A (charged or neutral) we may select as the reference that form of A which is of closed-shell character and differs from A by no more than 2 electrons. Then by using one of the EOM variants listed above we may recover an original structure. Owing to that we may avoid calculations based on the potentially spin-contaminated unrestricted HF reference. Moreover, the DIP and DEA approaches open the way to describe in a correct way a homolytic dissociation of the single bond without necessity to deal with the open-shell products. For all five considered methods (i.e., EE-EOM-CC, IP-EOM-CC, DIP-EOM-CC, EA-EOM-CC, DEA-EOM-CC) we provide detailed working equations both at the CCSD and CCSDT level in the form ready to code. For each of the considered EOM schemes we give illustrative results which make it possible to compare the performance of the E0M-CCSD and EOM-CCSDT approaches.

4.1 Introduction

Coupled-cluster (CC) [1–7] theory provides a powerful computational tool to study the electronic structure of atoms and molecules. As many of the Hartree–Fock (HF) based methods the CC wave function is expanded in terms of the reference and ”excited” configurations and the latter are

Quantum Chemistry and Dynamics of Excited States: Methods and Applications,
First Edition. Edited by Leticia González and Roland Lindh.

obtained by replacement of one-electron functions by their virtual counterparts in the reference single Slater determinant. The exponential parameterization of the wave function makes the CC method a robust computational scheme which in recent years has dominated the field of *ab initio* calculations. It is commonly known that one of the practical realizations of CC theory based on the inclusion of singles and doubles into the CC expansion augmented by non-iterative triples [8] is considered the ”gold” standard of quantum chemical methods and has tens of thousand of applications.

A principal realization of the CC approach is the method focused on properties and energetics of atoms and molecules with the ground state dominated by a single reference (SR). One of the possibilities to go beyond the SR formulation is to use its multi-reference extension (MR) [9]. The latter scheme offers a way to treat radicals and ionized species or, in general, systems which require a multi-determinantal description of the ground state. The MR-CC scheme also provides a possibility to describe excited states. However, due to a rather complicated formal background the MR-CC schemes have a moderate popularity and their best days are yet to come.

On the other hand applying CC theory to study atomic and molecular excited states has a long history which is connected with merging the CC approach with the equation-of-motion (EOM) formalism. The EOM approach goes back to the works of Rowe [10] who designed the general pattern of incorporating the time-dependent Schrödinger equation into the methodology of electronic structure calculations. Subsequent works [11–14] represent successful attempts to combine the EOM formalism with CC theory. However, the real breakthrough was achieved by Stanton and Bartlett who in their seminal paper [15] introduced the EOM-CCSD scheme into high level studies of excited electronic (EE) states. They derived the working equations in connection with the generalized Davidson diagonalization procedure, demonstrated the use of EOM-CCSD in property calculations and finally implemented the EOM-CCSD scheme into the ACES program package for general, i.e., UHF (Unrestricted HF) reference functions. From the same group, papers appeared later on extending the EOM formalism to ionization potential (IP) [16] calculations as well as to electron-attached (EA) [17] states. In the following years the EOM-CCSD scheme was extended to include also triply excited configurations, first partially [18–22] and then in a full manner [21–25]. A similar development, i.e., inclusion of full triples, was accomplished for the IP [21, 26–28] and EA [21, 28–30] variants.

In this chapter we will focus on the principal variants of the EOM-CC approach. Since this book is devoted to the description of excited states the obvious field of our study will be the EE-EOM-CC scheme. However, we would like to point out that the IP-EOM-CC scheme, developed to compute ionization potentials, can also be employed to study excited electronic states of singly ionized systems. Hence, e.g., to obtain the electronic states of the CN radical we may use either the UHF-based EE-EOM-CC method or the RHF-based IP-EOM-CC scheme applied to CN^- closed-shell ion for that purpose. Since closed-shell calculations are generally easier to carry out we will generally prefer the second option. The same strategy can be applied in the case of the EA-EOM-CC scheme which can be used to study electronic excitations in systems with an extra electron.

In the same spirit we may use the EOM-CC methods aimed at the calculation of the double ionization potential, i.e., the DIP-EOM-CC [31–35]. The latter scheme can be applied to study electronic states of a system A in the situations where the reference function is more easily recovered for the A^{2-} ion, i.e., for the system with two added electrons. Applying the DIP-EOM-CC method will provide excitation energies for the neutral A molecule. In an analogous way the EOM-CC approach can be applied to study systems with two attached electrons, DEA-EOM-CC (double-electron-attached EOM-CC) [35, 36]. Here, we compute the reference function for the A^{2+} ion and by doing DEA-EOM-CC calculations we again recover results for the neutral A species.

In the following we will describe in detail all five above mentioned realizations of the EOM-CC approach: EE-EOM-CC, IP-EOM-CC, EA-EOM-CC, DIP-EOM-CC and DEA-EOM-CC in the spin-free formulation (i.e., non-orthogonally spin-adapted) based on a closed-shell (restricted) reference. For each one we will consider two coupled-cluster models: CCSD and CCSDT supplemented with illustrative examples of the results.

In addition, EOM-CC is closely related to the coupled-cluster linear response (CCLR) method [37–41] whose eigenvalues are the same for the full models, e.g., CCSD, CCSDT etc., and to the SAC-CI (symmetry-adapted cluster configuration interaction) method [42–44] which is more like CI than CC with additional approximations.

4.2 Theoretical Background

In this section we will present a formal derivation of the EOM equations in connection with coupled-cluster theory. Our starting point is the Schrödinger equation for the ground (or more generally: reference) state:

$$\hat{H}|\Psi_0\rangle = E_0|\Psi_0\rangle \tag{4.1}$$

where the Hamiltonian $\hat{H}$ is expressed in the second-quantized formalism as:

$$\hat{H} = \sum_{pq} f_q^p \hat{a}_p^\dagger \hat{a}_q + \frac{1}{2}\sum_{pqrs} v_{rs}^{pq} \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r$$

It is more convenient to write the Schrödinger equation using the Hamiltonian operator in the normal-order form, $\hat{H}_N$:

$$\hat{H}_N|\Psi_0\rangle = \Delta E_0|\Psi_0\rangle \tag{4.2}$$

where $\Delta E_0 = E_0 - \langle\Phi_0|\hat{H}|\Phi_0\rangle$ and $|\Phi_0\rangle$ is a single Slater determinant, usually chosen to be the Hartree–Fock reference state and $\hat{H}_N$ is defined as:

$$\begin{aligned}\hat{H}_N &= \hat{H} - \langle\Phi_0|\hat{H}|\Phi_0\rangle \\ &= \sum_p e_p\{\hat{a}_p^\dagger \hat{a}_p\} + \sum_{p\neq q} f_q^p\{\hat{a}_p^\dagger \hat{a}_q\} + \frac{1}{2}\sum_{pqrs} v_{rs}^{pq}\{\hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r\} \\ &= \hat{H}_N^o + \hat{F}_N + \hat{V}_N\end{aligned} \tag{4.3}$$

where e_p are one-particle eigenenergies and $\hat{F}_N$ and $\hat{V}_N$ are one- and two-body operators with $\hat{F}_N$ disappearing for the canonical Hartree–Fock reference state and v_{rs}^{pq} is a two-electron integral defined as:

$$v_{rs}^{pq} = \iint \varphi_p^*(1)\varphi_q^*(2)\frac{1}{r_{12}}\varphi_r(1)\varphi_s(2)dv_1 dv_2 \tag{4.4}$$

This v_{rs}^{pq} symbol introduces a more compact form of the integrals compared to that given in Eq. (1.4) of the Introductory Chapter which makes them more convenient for the cumbersome EOM-CC equations. The braces $\{\hat{a}_p\hat{a}_q\}$ indicate the normal-order form of the operators $\hat{a}_p\hat{a}_q$.

4.2.1 Coupled-Cluster Wave Function

The ground state wave function within the coupled-cluster [1–8, 45–54] formalism is expressed as:

$$|\Psi_0\rangle = e^{\hat{T}}|\Phi_0\rangle \tag{4.5}$$

where $\hat{T}$ is a cluster operator responsible for excitations:

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + ... + \hat{T}_N = \sum_{a,i} t_i^a \hat{a}_a^\dagger \hat{a}_i + \frac{1}{2} \sum_{ab,ij} t_{ij}^{ab} \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_j \hat{a}_i + \frac{1}{6} \sum_{abc,ijk} t_{ijk}^{abc} \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_c^\dagger \hat{a}_k \hat{a}_j \hat{a}_i + ... \quad (4.6)$$

N is the number of electrons in the system and the coefficients $t_{ij...}^{ab\cdots}$ are cluster amplitudes. Note that $\hat{a}_i, \hat{a}_j, \cdots$ second-quantized operators which annihilate electrons from occupied levels become – in the particle–hole formalism – creation operators (responsible for creation of holes). In connection with that the second-quantized operators present in the definition of the $\hat{T}$ operators are all creation operators. The same applies to the $\hat{R}$ operators defined in the following sections.

Inserting the wave function into the Schrödinger equation

$$\hat{H}_N e^{\hat{T}} |\Phi_0\rangle = \Delta E_0 e^{\hat{T}} |\Phi_0\rangle \quad (4.7)$$

we obtain the desired solutions, i.e., by multiplying from the left with $e^{-\hat{T}}$ and projecting against $|\Phi_0\rangle$ we get the correlation energy:

$$\Delta E_0 = \langle \Phi_0 | e^{-T} \hat{H}_N e^T | \Phi_0 \rangle = \langle \Phi_0 | \hat{\bar{H}}_N | \Phi_0 \rangle \quad (4.8)$$

Projecting against the configurations $|\Phi_{ij...}^{ab\cdots}\rangle$ we obtain the coupled-cluster equations:

$$\langle \Phi_{ij...}^{ab\cdots} | \hat{\bar{H}}_N | \Phi_0 \rangle = 0 \quad (4.9)$$

In the equations above we introduced the $\hat{\bar{H}}_N$ symbol representing a similarity transformed Hamiltonian:

$$\hat{\bar{H}}_N = e^{-\hat{T}} \hat{H}_N e^{\hat{T}} \quad (4.10)$$

The similarity-transformed Hamiltonian $\hat{\bar{H}}_N$ is a crucial quantity in the derivation and exploitation of the EOM scheme. Due to the Baker–Campbell–Hausdorff formula:

$$\hat{\bar{H}}_N = \hat{H}_N + [\hat{H}_N, \hat{T}] + \frac{1}{2}[[\hat{H}_N, \hat{T}], \hat{T}] + \frac{1}{6}[[[\hat{H}_N, \hat{T}], \hat{T}], \hat{T}] + \frac{1}{24}[[[[\hat{H}_N, \hat{T}], \hat{T}], \hat{T}], \hat{T}]$$

$\hat{\bar{H}}_N$ is expressed as a terminating commutator series, hence all CC equations occur in closed form.

4.2.2 The Equation-of-Motion Approach

The equation-of-motion [7, 11–36, 55–63] approach is a useful tool for the study of states obtained via excitation, ionization and electron attachment processes for which the ground state is described within CC theory. To begin with we do not need to specify the type of process involved. We assume that the wave function $|\Psi_k\rangle$ resulting from the action of the linear operator $\hat{R}(k)$ on the ground state wave function $|\Psi_0\rangle$:

$$|\Psi_k\rangle = \hat{R}(k) |\Psi_0\rangle \quad (4.11)$$

is a solution of the Schrödinger equation:

$$\hat{H}_N |\Psi_k\rangle = \Delta E_k |\Psi_k\rangle \quad (4.12)$$

or

$$\hat{H}_N \hat{R}(k) |\Psi_0\rangle = \Delta E_k \hat{R}(k) |\Psi_0\rangle \quad (4.13)$$

where the $\Delta E_k = E_k - \langle \Phi_0 | H | \Phi_0 \rangle$ and E_k is the total energy of the target k-state. Multiplying Eq. (4.2) by $\hat{R}(k)$ from the left and subtracting from the above we obtain:

$$[\hat{H}_N, \hat{R}(k)] |\Psi_0\rangle = \omega_k \hat{R}(k) |\Psi_0\rangle \quad (4.14)$$

where $\omega_k = \Delta E_k - \Delta E_0 = E_k - E_0$ is the transition energy connected with the studied process. Substituting the exponential ansatz, Eq. (4.5), for $|\Psi_0\rangle$ and multiplying from the left with $e^{-\hat{T}}$ we obtain:

$$e^{-\hat{T}}[\hat{H}_N, \hat{R}(k)]e^{\hat{T}}|\Phi_0\rangle = \omega_k e^{-\hat{T}}\hat{R}(k)e^{\hat{T}}|\Phi_0\rangle \tag{4.15}$$

Since the operators $\hat{T}$ and $\hat{R}$ commute, we finally get the EOM-CC equation:

$$[\hat{H}_N, \hat{R}(k)]|\Phi_0\rangle = \omega_k \hat{R}(k)|\Phi_0\rangle \tag{4.16}$$

Defining the normal-ordered similarity-transformed Hamiltonian in a slightly different manner

$$\bar{H}_N = e^{-\hat{T}}\hat{H}e^{\hat{T}} - \langle\Phi_0|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Phi_0\rangle \tag{4.17}$$

we may write the EOM equation as an eigenproblem of the $\bar{H}_N$ operator:

$$\bar{H}_N\hat{R}(k)|\Phi_0\rangle = \omega_k\hat{R}(k)|\Phi_0\rangle \tag{4.18}$$

The most general form of the $\hat{R}(k)$ operator can be written out as:

$$\hat{R}(k) = r_0(k) + \hat{R}_1(k) + \hat{R}_2(k) + \ldots + \hat{R}_N(k) \tag{4.19}$$

Some of the components of $\hat{R}(k)$ disappear for specific processes, e.g., r_0 vanishes for ionization and electron attachment, $\hat{R}_1$ does not occur for double ionization and double electron attachment, etc.

Note that the $\bar{H}_N$ operator is not Hermitian, hence we will have two types of eigenvectors: the right-hand solution, Eq. (4.18), and the left-hand solution:

$$\langle\Phi_0|\hat{L}(k)\bar{H}_N = \langle\Phi_0|\hat{L}(k)\omega_k \tag{4.20}$$

$\hat{L}$ is a de-excitation operator and the two sets of eigenfunctions $\hat{R}|\Phi_0\rangle$ and $\langle\Phi_0|\hat{L}$ are bi-orthonormal, i.e.,

$$\langle\Phi_0|\hat{L}(k)\hat{R}(l)|\Phi_0\rangle = \delta_{kl} \tag{4.21}$$

Both $\hat{L}(k)$ and $\hat{R}(k)$ have the same ω_k eigenvalue and both are needed to obtain ordinary and transition density matrices but both bra and ket states are not needed if we are only interested in the energy eigenvalues.

4.2.3 Similarity-Transformed Hamiltonian

The main step in EOM calculations is the diagonalization of the $\bar{H}_N$ operator within the proper configurational subspace. The type of subspace depends on the class of the EOM problem under consideration. For example, solving the EE-EOM-CCSD problem, we diagonalize the $\bar{H}_N$ operator in the subspace of $|\Phi_i^a\rangle$ and $|\Phi_{ij}^{ab}\rangle$ configurations; going to the CCSDT model we expand the working space to include triple excitation configurations $|\Phi_{ijk}^{abc}\rangle$. Performing IP-EOM-CCSD calculations we limit the diagonalization subspace to the configurations $|\Phi_i\rangle$ and $|\Phi_{ij}^a\rangle$, i.e., representing the single ionization plus those with an additional single excitation. Going to the IP-EOM-CCSDT model we add the $|\Phi_{ijk}^{ab}\rangle$ configurations. Analogous rules apply to all other classes of EOM problems. For EA-EOM-CC the respective subspace includes $|\Phi^a\rangle$, $|\Phi_i^{ab}\rangle$ and $|\Phi_{ij}^{abc}\rangle$ (CCSDT) configurations; for DIP-EOM-CC: $|\Phi_{ij}\rangle$ and $|\Phi_{ijk}^a\rangle$ (CCSDT); for DEA-EOM-CC $|\Phi^{ab}\rangle$ and $|\Phi_i^{abc}\rangle$ (CCSDT).

An additional comment is needed with respect to the structure of the EOM matrix in case of the treatment of electronic excitations. The $\bar{\mathbf{H}}_\mathbf{N}$ matrix corresponding to the CCSD model can be expressed in the following way, where S represents the $|\Phi_i^a\rangle$ and D – the $|\Phi_{ij}^{ab}\rangle$ configurations:

$$\bar{\mathbf{H}}_\mathbf{N} = \begin{bmatrix} \langle 0|\bar{H}_N|0\rangle & \langle 0|\bar{H}_N|S\rangle & \langle 0|\bar{H}_N|D\rangle \\ 0 & \langle S|\bar{H}_N|S\rangle & \langle S|\bar{H}_N|D\rangle \\ 0 & \langle D|\bar{H}_N|S\rangle & \langle D|\bar{H}_N|D\rangle \end{bmatrix}$$

The matrix elements $\langle\Phi_i^a|\bar{H}_N|\Phi_0\rangle$ ($\equiv\langle S|\bar{H}_N|0\rangle$) and $\langle\Phi_{ij}^{ab}|\bar{H}_N|\Phi_0\rangle$ ($\equiv\langle D|\bar{H}_N|0\rangle$) are 0 due to the fulfillment of the CC equations, Eq. (4.9). Hence, diagonalization of $\bar{H}_N$ within the subspace of the excited configurations will give the same eigenvalues as a diagonalization of the full matrix. In the latter case, we also get the $\langle 0|\bar{H}_N|0\rangle$ ($\equiv\langle\Phi_0|\bar{H}_N|\Phi_0\rangle = 0$) reference energy unchanged.

The full EOM-CCSDT method requires solution of the CC equations, Eq. (4.9), for the T_1, T_2 and T_3 amplitudes and then construction of the $\bar{H}_N$ operator up to four-body elements, according to Eq. (4.17). In column 1 of Table 4.1 we list all categories of the $\bar{H}_N$ operator and in column 2 the terms contributing to each $\bar{H}_N$ element.

Emphasizing the many-body structure of the $\bar{H}_N$ we may decompose it into individual n-body contributions I^n as follows:

$$\bar{H}_N = I^1 + I^2 + I^3 + I^4 \tag{4.22}$$

We skipped the $\bar{H}_N$ components with 0 annihilation lines since they cannot be contracted with the $\hat{R}$ operators. We can easily identify a particular term in column 1 of Table 4.1 with the proper I^n component of Eq. (4.22) bearing in mind that the term with the total number of indices equal to $2n$ belongs to the I^n element of $\bar{H}_N$. In Table 4.1 the one-, two-, three- and four-body terms are separated with solid horizontal lines. We may further classify the particular I^n term into I_k^n components where the subscript k indicates the number of annihilation lines (terminology used in the diagrammatic formulation of the EOM equations [7]) or – equivalently – the number of particle–hole annihilation operators.

$$\begin{aligned} I^1 &= I_1^1 + I_2^1 \\ I^2 &= I_1^2 + I_2^2 + I_3^2 + I_4^2 \\ I^3 &= I_1^3 + I_2^3 + I_3^3 \\ I^4 &= I_2^4 + I_3^4 \end{aligned} \tag{4.23}$$

The number of particle–hole annihilation operators in each term of column 1 is equal to the sum of particle indices in the subscript position and hole indices in the superscript position. E.g., the first two-body term in Table 4.1, I_{bc}^{ai} is connected with three annihilation operators: one (i) hole-index as a superscript and two (b, c) particle-indices as subscripts. Similarly, the first three-body I-component, I_{jkl}^{iab} is connected with one annihilation operator (one annihilation line in the diagrammatic formalism) since there is only one hole-index (i) in the superscript position and zero particle-indices in the subscript position. The presence of annihilation operators is crucial in the derivation of the EOM equations since only those can be contracted with the R operators and the number of summation indices in the terms contributing to the EOM equation is equal to the number of annihilation lines in the respective $\bar{H}_N$ element. Note that not all four-body terms are listed in Table 4.1 but only those which occur in the EOM-CC equations up to the CCSDT level. This is why we skipped the I_1^4 element in Eq. (4.23) in the EOM-CCSDT model. Moreover, throughout the chapter we adopted tensor notation with implied summation over repeated indices.

4.2.4 Davidson Diagonalization Algorithm

In this section we indicate the main points of the diagonalization of large matrices developed by Davidson [64] which is a crucial procedure in the solution of the EOM-CC equations. The matrix diagonalization is also an essential step in the many quantum chemical schemes, compare the MRPT approach presented in Chapter 2. The original work of Davidson applied to Hermitian

Table 4.1 Algebraic expressions for the elements of $\bar{H}_N$ used in the EOM-CCSD and EOM-CCSDT models (orbital formalism).

	Expression[a]
I_a^i	$f_a^i + 2t_n^e v_{ae}^{in} - t_n^e v_{ea}^{in}$
I_b^a	$(1-\delta_b^a)f_b^a + 2t_n^e v_{be}^{an} - t_n^e v_{be}^{na} - 2t_{no}^{ea} v_{eb}^{no} + t_{no}^{ea} v_{be}^{no} - t_n^a I_b^n$
I_j^i	$(1-\delta_j^i)f_j^i + 2t_n^e v_{je}^{in} - t_n^e v_{ej}^{in} + 2t_{nj}^{ef} v_{ef}^{ni} - t_{nj}^{ef} v_{ef}^{in} + t_j^e I_e^i$
I_{bc}^{ai}	$I'^{ai}_{bc} - \frac{1}{2} t_n^a v_{bc}^{ni}$
I'^{ai}_{bc}	$v_{bc}^{ai} - \frac{1}{2} t_n^a v_{bc}^{ni}$
I_{ka}^{ij}	$I'^{ij}_{ka} + \frac{1}{2} t_k^e v_{ea}^{ij}$
I'^{ij}_{ka}	$v_{ka}^{ij} + \frac{1}{2} t_k^e v_{ea}^{ij}$
I_{cd}^{ab}	$v_{cd}^{ab} - P(ac/bd) t_n^b I'^{an}_{cd} + t_{no}^{ab} v_{cd}^{no}$
I_{kl}^{ij}	$v_{kl}^{ij} + P(ik/jl) t_l^e I'^{ij}_{ke} + t_{kl}^{ef} v_{ef}^{ij}$
I_{bi}^{aj}	$I''^{aj}_{bi} - t_{ni}^{ae} v_{be}^{nj} + \frac{1}{2} t_i^e I_{be}^{aj}$
I'^{aj}_{bi}	$v_{bi}^{aj} - \frac{1}{2} t_n^a v_{bi}^{nj} + t_i^e I'^{aj}_{be}$
I''^{aj}_{bi}	$v_{bi}^{aj} - t_n^a v_{bi}^{nj} + \frac{1}{2} t_i^e I_{be}^{aj}$
I_{ib}^{aj}	$I''^{aj}_{ib} + 2t_{ni}^{ea} v_{eb}^{nj} - t_{in}^{ae} v_{eb}^{jn} - t_{ni}^{ae} v_{eb}^{nj} + \frac{1}{2} t_i^e I_{eb}^{aj}$
I'^{aj}_{ib}	$v_{ib}^{aj} - \frac{1}{2} t_n^a v_{ib}^{nj} + t_i^e I'^{aj}_{eb}$
I''^{aj}_{ib}	$v_{ib}^{aj} - t_n^a v_{ib}^{nj} + \frac{1}{2} t_i^e I_{eb}^{aj}$
I_{ci}^{ab}	$I'^{ab}_{ci} - 2t_{noi}^{aeb} v_{ce}^{no} + t_{nio}^{aeb} v_{ce}^{no} + t_{noi}^{aeb} v_{ec}^{no}$
I'^{ab}_{ci}	$v_{ci}^{ab} + t_i^e v_{ce}^{ab} - t_n^b I'^{an}_{ci} - t_n^a I'^{nb}_{ci} - t_{ni}^{ab} I_c^n + 2t_{ni}^{eb} I_{ce}^{an} - t_{ni}^{eb} I_{ce}^{na} - t_{ni}^{ae} I_{ce}^{nb} + t_{no}^{ab} I_{ci}^{no}$
I_{jk}^{ia}	$I'^{ia}_{jk} + 2t_{jnk}^{efa} v_{ef}^{in} - t_{jkn}^{efa} v_{ef}^{in} - t_{jnk}^{efa} v_{ef}^{ni}$
I'^{ia}_{jk}	$v_{jk}^{ia} - t_n^a v_{jk}^{in} + t_k^e I''^{ia}_{je} + t_j^e I''^{ia}_{ek} + t_{jk}^{ea} I_e^i + 2t_{nk}^{ea} I_{je}^{in} - t_{nk}^{ea} I_{ej}^{in} - t_{jn}^{ea} I_{ek}^{in} + t_{jk}^{ef} I_{ef}^{ia}$
I_{jkl}^{iab}	$I'^{iab}_{jkl} + P(ka/lb) t_{jkl}^{efb} I_{ef}^{ia}$
I'^{iab}_{jkl}	$-P(ka/lb)[t_{nl}^{ab} I_{jk}^{in} - t_{kl}^{eb} I_{je}^{ia} - t_{jl}^{eb} I_{ek}^{ia}] + t_{jkl}^{eab} I_e^i - P(ka/lb)[t_{jnl}^{eab} I_{ek}^{in} + t_{nkl}^{aeb} I_{je}^{in}] + 2t_{nkl}^{eab} I_{je}^{in} - t_{nkl}^{eab} I_{ej}^{in}$
I_{dij}^{abc}	$-P(ib/jc)[t_{mj}^{bc} I_{di}^{am} + t_{mj}^{ac} I_{di}^{mb} - t_{ij}^{fc} I_{df}^{ab} + t_{mij}^{afc} I_{df}^{mb} + t_{mij}^{bfc} I_{df}^{am} - t_{mnj}^{abc} I_{di}^{mn}] + 2t_{mij}^{fbc} I_{df}^{am} - t_{mij}^{fbc} I_{df}^{ma} + t_{mij}^{abc} I_d^m$
I_{klm}^{ija}	$P(ik/jl) t_{lm}^{ea} I_{ke}^{ij} + t_{klm}^{efa} v_{ef}^{ij}$
I_{cjk}^{aib}	$I'^{aib}_{cjk} - t_{njk}^{aeb} v_{ce}^{ni}$
I'^{aib}_{cjk}	$t_{jk}^{eb} I_{ce}^{ai} - t_{nk}^{ab} I_{cj}^{ni}$
I_{cjk}^{iab}	$I'^{iab}_{cjk} + 2t_{njk}^{eab} v_{ce}^{in} - P(ja/kb) t_{jnk}^{eab} v_{ce}^{in} - t_{njk}^{eab} v_{ce}^{ni}$
I'^{iab}_{cjk}	$P(ja/kb)[t_{jk}^{eb} I_{ce}^{ia} - t_{nk}^{ab} I_{cj}^{in}]$
I_{ilb}^{ajk}	$t_{il}^{ae} v_{eb}^{jk}$
I_{icd}^{abj}	$-t_{in}^{ab} v_{cd}^{nj}$
I_{dei}^{abc}	$-P(ad/be) t_{mi}^{bc} I_{de}^{am} + t_{mni}^{abc} v_{de}^{mn}$
I_{jbkl}^{iacd}	$-t_{mkl}^{acd} I_{jb}^{im} + t_{jkl}^{ecd} I_{be}^{ai} - P(kc/ld) t_{km}^{ca} I_{bjl}^{mid}$
I_{ajkl}^{ibcd}	$-t_{mkl}^{bcd} I_{aj}^{im} + t_{jkl}^{ecd} I_{ae}^{ib} - P(kc/ld) t_{ml}^{cd} I_{jak}^{bim}$
I_{aklm}^{ijbc}	$t_{klm}^{ebc} v_{ae}^{ij}$
I_{aejk}^{ibcd}	$-t_{mjk}^{bcd} v_{ae}^{im}$

a) $P(ka/lb)$ implies sum of the additional components with ka permuted with lb and similarly to the rest of $P(pq/rs)$. δ_q^p designates the usual Kronecker delta.

matrices was later generalized to non-Hermitian cases by Nakatsuji and Hirao [65]. The Davidson procedure is aimed at recovering only a few eigenvalues while the dimension of the matrix, in this case it is $\bar{\mathbf{H}}_\mathbf{N}$, may go to millions and larger dimensions. We assume the initial (guess) vector **r1** which is in general a very crude approximation to the true eigenvector of $\bar{\mathbf{H}}_\mathbf{N}$. Taking the product

$$\mathbf{b_1} = \bar{\mathbf{H}}_\mathbf{N} \times \mathbf{r1}$$

and the dot product $\lambda_1 = \langle \mathbf{b_1} | \mathbf{r_1} \rangle$ we get an approximation to the $\bar{\mathbf{H}}_\mathbf{N}$ eigenvalue, in general very far from the true value. Then we construct the error vector **e**

$$\mathbf{e} = \mathbf{b_1} - \lambda_1 \mathbf{r1}$$

and on the basis of the latter we obtain a vector **r2** (called the Davidson correction vector) which initiates the second iteration and the elements of which are obtained as:

$$r2_i = e_i/(\lambda_1 - \bar{H}_{ii}) \;\; i = 1, 2, ..., n$$

where $\bar{H}_{ii}$ is a diagonal element of the $\bar{\mathbf{H}}_\mathbf{N}$ matrix. Next we take the product $\bar{\mathbf{H}}_\mathbf{N} \times r2$ to obtain the $\mathbf{b_2}$ vector and the situation is repeated. In each iteration the crucial step is taking the product of $\bar{\mathbf{H}}_\mathbf{N}$ and the $\mathbf{r_i}$ vector. This is the most time consuming and difficult step to carry out. In order to avoid a construction of the $\bar{\mathbf{H}}_\mathbf{N}$ matrix we take the product $\bar{H} \times \hat{R}$ at the operator level. Applying Wick's theorem and performing proper contractions we obtain elements of the target vector $\bar{\mathbf{H}}_\mathbf{N}\mathbf{R}$. The EOM equations which are presented in the next subsection represent contributions to the target $\bar{\mathbf{H}}_\mathbf{N}\mathbf{R}$ vector via contractions of the $\bar{H}$ component $I_{rs\ldots}^{pq\cdots}$ and the elements of the $\hat{R}$ operator. Thus using the $\bar{H}$ and $\hat{R}$ components we can obtain the elements of the target vector without construction of the large $\bar{\mathbf{H}}_\mathbf{N}$ matrix.

4.3 Excited States: EE-EOM-CC

From the analysis presented above it follows that the crucial step in a real calculation based on the generalized Davidson algorithm is a construction of the product of $(\bar{H}_N R)$. In most cases this product is calculated by taking contractions between the $\bar{H}_N$ and $\hat{R}$ operators. As was shown in the previous subsection some of the $\bar{H}_N$ elements (e.g., three-body and four-body ones) are difficult and costly to compute and storing them would require a large amount of disk space. In order to make the EOM scheme more feasible some contributions to $(\bar{H}_N R)$ are calculated according to the original expression:

$$\bar{H}_N \hat{R} = e^{-\hat{T}} \hat{H}_{\tilde{N}} e^{\hat{T}} \hat{R} = e^{-\hat{T}} \hat{H}_{\tilde{N}} \hat{R} e^{\hat{T}} = e^{-\hat{T}} \chi e^{\hat{T}} \tag{4.24}$$

where $\hat{H}_{\tilde{N}} = \hat{H} - \langle \Psi_0 | \hat{H} | \Psi_0 \rangle$ and the second equality results from the fact that $\hat{R}$ and $e^{\hat{T}}$ operators commute. So in order to avoid cumbersome construction of the $\bar{H}_N$ element we may first contract $\hat{H}_N$ and $\hat{R}$ to get the χ intermediate. Then, to get the contribution to $(\bar{H}_N R)$ we need to make the $\chi e^{\hat{T}}$ contraction. However, the price we pay for the construction of the simpler χ intermediate is the fact that we have to do it in each iteration while the $\bar{H}_N$ elements are computed only once. Below we present the detailed contributions to the $(\bar{H}_N R)$ elements both at the CCSD and CCSDT levels.

4.3.1 EE-EOM-CCSD Model

In the singles and doubles approximation (EOM-CCSD) the $\hat{R}(k)$ operator takes the form:

$$\hat{R}(k)_{SD} = r_0(k) + \sum_a^{N_v} \sum_i^{N_o} r_i^a(k) \hat{a}_a^\dagger \hat{a}_i + \frac{1}{2} \sum_{ab}^{N_v} \sum_{ij}^{N_o} r_{ij}^{ab}(k) \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_j \hat{a}_i \tag{4.25}$$

Table 4.2 The EE-EOM-CCSD equations in the orbital formalism.

	Expression[a)]
$(\bar{H}_N R)_i^a =$	$-r_m^a I_i^m + r_i^e I_e^a - r_m^e I_{ie}^{ma} + 2r_m^e I_{ei}^{ma} + 2r_{im}^{ae} I_e^m$
	$-r_{im}^{ea} I_e^m - 2r_{mn}^{ae} I_{ie}^{mn} + r_{mn}^{ea} I_{ie}^{mn} + 2r_{im}^{ef} I_{ef}^{am} - r_{mi}^{ef} I_{ef}^{am}$
$(\bar{H}_N R)_{ij}^{ab} =$	$P(ia/jb)[r_i^e I_{ej}^{\prime ab} - r_m^a I_{ij}^{\prime mb} - r_{im}^{ab} I_j^m + r_{ij}^{ae} I_e^b$
	$-r_{mj}^{eb} I_{ie}^{ma} - r_{mj}^{ae} I_{ie}^{mb} + 2r_{mj}^{eb} I_{ei}^{ma} - r_{im}^{ea} I_{ej}^{mb}]$
	$+r_{mn}^{ab} I_{ij}^{mn} + r_{ij}^{ef} I_{ef}^{ab} + F_{ij}^{ab}$
$F_{ij}^{ab\,b)} =$	$P(ia/jb)[-r_m^e I_{iej}^{\prime mab} - 2r_{mn}^{ae} I_{iej}^{mnb} - r_{mi}^{ef} I_{efj}^{amb}$
	$+2r_{im}^{ef} I_{efj}^{amb} + r_{mn}^{ea} I_{iej}^{mnb}] + 2r_m^e I_{eij}^{\prime mab}$
$F_{ij}^{ab\,c)} =$	$P(ia/jb)[t_{ij}^{ae} \chi_e^b - t_{im}^{ab} \chi_j^{\prime m}]$
$\chi_j^{\prime i} =$	$2r_m^e I_{je}^{im} - r_m^e I_{je}^{mi} + 2r_{jm}^{ef} v_{ef}^{im} - r_{jm}^{ef} v_{ef}^{mi}$
$\chi_b^a =$	$2r_m^e I_{be}^{am} - r_m^e I_{eb}^{am} - 2r_{mn}^{ae} v_{be}^{mn} + r_{mn}^{ae} v_{eb}^{mn}$

a) $P(ia/jb)$ implies sum of the additional components with ia permuted with jb; all elements of $\bar{H}_N$ ($I_{rs\ldots}^{pq\ldots}$) are defined in Table 4.1.
b) Standard version.
c) Factorized version.

where $r_0(k)$ is a non-zero constant if the kth state has the same symmetry as the ground state, otherwise is zero. N_o (N_v) indicates the number of occupied (virtual) levels in the system. The $\hat{R}(k)$ operator can be expressed through the elementary creation-annihilation operators. It is obvious that by operating with the sequence $\hat{a}_a^\dagger \hat{a}_i$ on the Fermi vacuum $|\Phi_0\rangle$ we obtain a singly excited configuration:

$$\hat{a}_a^\dagger \hat{a}_i |\Phi_0\rangle = |\Phi_i^a\rangle$$

The algebraic contributions to the $\bar{H}_N R$ elements within the considered model are shown in Table 4.2. The adoption of the orbital formalism means that we have integrated over spin components (hence in some terms we have the factor of 2) and resulting equations are spin-free.

Before presenting the explicit formula for the target ($\bar{H}_N R$) amplitudes we return to equation (4.18) to evaluate the r_0 parameter. Solving the EOM-CC equation (4.18) we get the ω eigenvalue which can be used to determine the r_0 parameter. Projecting Eq. (4.18) against $|\Phi_0\rangle$ and bearing in mind the expression for $\hat{R}(k)$, Eq. (4.25), we obtain:

$$\omega_k r_0(k) = \langle \Phi_0 | \bar{H}_N (\hat{R}_1(k) + \hat{R}_2(k)) | \Phi_0 \rangle$$

which can be transformed into the working formula:

$$r_0(k) = (2 \sum_{ia} I_a^i r_i^a(k) + \sum_{ijab} (2v_{ab}^{ij} - v_{ab}^{ji}) r_{ij}^{ab}(k)) / \omega_k$$

Note that to each term in Tables 4.2 and 4.3 the proper permutation of the external indices should be applied.

We may also illustrate here how the factorization works in practice. In Table 4.2 we have the quantity F_{ij}^{ab} marked with b) contributing to the $(\bar{H}_N R)_{ij}^{ab}$ element which is computed via the standard formula based on the three-body elements of $\bar{H}_N$ requiring in most cases an n^7 scaling (here

Table 4.3 The EE-EOM-CCSDT equations in orbital formalism (additional terms not included in the EE-EOM-CCSD model).

	Expression[a)]
$(\bar{H}_N R)_i^a =$	$2r_{mni}^{efa} v_{ef}^{mn} - 2r_{min}^{efa} v_{ef}^{mn} - r_{mni}^{fea} v_{ef}^{mn} + r_{imn}^{efa} v_{ef}^{mn}$
$(\bar{H}_N R)_{ij}^{ab} =$	$P(ia/jb)[r_i^e I_{ej}^{ab} - r_m^a I_{ij}^{mb} + 2r_{mij}^{efb} I_{ef}^{ma} - r_{imj}^{efb} I_{ef}^{ma} - r_{imj}^{eaf} I_{ef}^{mb}$
	$-2r_{mnj}^{eab} I_{ei}^{mn} + r_{mnj}^{aeb} I_{ei}^{mn} + r_{imn}^{eab} I_{ej}^{mn} - 1r_{imj}^{eab} I_e^m] + 2r_{mij}^{eab} I_e^m + F_{ij}^{ab}$
$F_{ij}^{ab\,b)} =$	$-P(ia/jb) - r_m^e I_{iej}^{mab} + 2r_m^e I_{eij}^{mab}$
$F_{ij}^{ab\,c)} =$	$-P(ia/jb) t_{imj}^{eab} \chi_e^m + 2t_{mij}^{eab} \chi_e^m$
$(\bar{H}_N R)_{ijk}^{abc} =$	$P(ia/jb\ kc)[-r_m^a I_{ijk}^{mbc} + r_{mn}^{bc} I_{ijk}^{amn} + r_{ijk}^{ebc} I_e^a - r_{mjk}^{abc} I_i^m$
	$+r_{ijk}^{aef} I_{ef}^{bc} + r_{imn}^{abc} I_{jk}^{mn} - r_{mjk}^{ebc} I_{ie}^{ma} + 2r_{mjk}^{ebc} I_{ei}^{ma}]$
	$+P(ia/jb/kc)[r_{ij}^{ae} I_{ek}^{bc} - r_{im}^{ab} I_{jk}^{mc} - r_{imk}^{abe} I_{je}^{mc} - r_{imk}^{eac} I_{ej}^{mb}] + F_{ijk}^{abc}$
$F_{ijk}^{abc\ \ b)} =$	$P(ia/jb\ kc)[r_i^e I_{ejk}^{abc} - r_m^e I_{iejk}^{mabc} + 2r_{mi}^{ea} I_{ejk}^{mbc} - r_{im}^{ea} I_{ejk}^{mbc}$
	$+r_{jk}^{ef} I_{ief}^{abc} + 2r_{im}^{ef} I_{efjk}^{ambc} - r_{mi}^{ef} I_{efjk}^{ambc} - 2r_{mn}^{ae} I_{iejk}^{mnbc} + r_{mn}^{ea} I_{iejk}^{mnbc}]$
	$+2r_m^e I_{eijk}^{mabc} + P(ia/jb/kc)[-r_{mk}^{ec} I_{iej}^{mab} - r_{mj}^{ae} I_{iek}^{mbc} + 2r_{mij}^{efb} I_{efk}^{mac}$
	$-r_{mij}^{efb} I_{efk}^{amc} - r_{imj}^{eaf} I_{efk}^{mbc} - 2r_{mnj}^{eab} I_{eik}^{mnc} + r_{mnj}^{eab} I_{iek}^{mnc} + r_{imn}^{eab} I_{ejk}^{mnc}]$
$F_{ijk}^{abc\ \ c)} =$	$P(ia/jb\ kc)[-t_{mjk}^{abc} \chi_i^m + t_{ijk}^{ebc} \chi_e^a - t_{mjk}^{ebc} \chi_{ei}^{am} + 2t_{mjk}^{ebc} I \chi_{ei}^{ma}$
	$+t_{imn}^{abc} \chi_{jk}^{mn}] + P(ia/jb/kc)[-t_{im}^{ab} \chi_{jk}^{mc} + t_{ij}^{ae} \chi_{ek}^{bc} - t_{imk}^{ebc} \chi_{ej}^{am} - t_{imj}^{eab} \chi_{ek}^{mc}]$

a) $P(ia/jb)$ implies sum of the additional components with ia permuted with jb; the symmetrizers $P(ia/jb/kc)$ and $P(ia/jb\ kc)$ are defined in a similar way to $P(ia/jb)$; all elements of $\bar{H}_N$ ($I_{rs\ldots}^{pq\cdots}$) are defined in Table 4.1 whereas the intermediates $\chi_{q\ldots}^{p\cdots}$ are defined in Tables 4.2 and 4.4.
b) Standard version.
c) Factorized version.

and elsewhere in the chapter n is the number of correlated occupied and unoccupied orbitals, i.e., $n = N_o + N_v$). The same contribution can be obtained via F_{ij}^{ab} marked with c), computed with an n^5 scaling which engages the χ_e^b and χ_j^m intermediates. The formulas for the latter quantities are given in the same table and, as we can see, they are also obtained with an n^5 scaling. It should be clarified though that since they engage the contraction with the $\hat{R}$ operator, they must be computed in each iteration.

4.3.2 EE-EOM-CCSDT Model

Within the EOM-CCSDT model the $\hat{R}(k)_{SDT}$ operator includes also the component responsible for triple excitations:

$$\hat{R}(k)_{SDT} = \hat{R}(k)_{SD} + \hat{R}_3(k) = \hat{R}(k)_{SD} + \frac{1}{6}\sum_{abc}^{N_v}\sum_{ijl}^{N_o} r_{ijl}^{abc}(k)\hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_c^\dagger \hat{a}_l \hat{a}_j \hat{a}_i$$

Inclusion of triply excited configurations is connected with much more complicated formulas for the $\bar{H}_N R$ elements and with much larger computational effort due to larger scaling, n^8 vs. n^6 (in case of EOM-CCSD).

The expressions listed in Table 4.3 include terms additional to Table 4.2 which are due to the $\hat{T}_3$ and $\hat{R}_3$ operators. Thus we observe the contributions to $(\bar{H}_N R)_i^a$ to include some terms originating

Table 4.4 Algebraic expression for the $\chi^{p\ldots}_{q\ldots}$ intermediates used inthe EE-EOM-CCSDT model in orbital formalism (additional terms not included in the EE-EOM-CCSD model).

	Expression[a)]
$\chi^i_j =$	$\chi'^i_j + r^e_j I^i_e$
$\chi^i_a =$	$-r^e_m v^{mi}_{ae} + 2r^e_m v^{mi}_{ea}$
$\chi^{ia}_{jk} =$	$2r^{ae}_{km} I^{mi}_{aj} - r^{ea}_{km} I^{mi}_{ej} - r^{ea}_{jm} I^{im}_{ek} - r^{ae}_{km} I^{mi}_{je} + t^{ea}_{jk} \chi^i_e + r^{ef}_{jk} I^{ia}_{ef}$
	$+2r^{aef}_{kmj} v^{mi}_{ef} - r^{aef}_{kjm} v^{mi}_{ef} - r^{eaf}_{kmj} v^{mi}_{ef} + r^e_j I^{ia}_{ek} + r^e_k I^{ia}_{je}$
$\chi^{ab}_{ci} =$	$2r^{be}_{im} I^{ma}_{ec} - r^{eb}_{im} I^{ma}_{ec} - r^{ae}_{mi} I^{mb}_{ce} - r^{be}_{im} I^{ma}_{ce} + r^e_i I^{ba}_{ec}$
$\chi^{ij}_{kl} =$	$r^{ef}_{kl} I^{ij}_{ef} + r^e_k I^{ij}_{el}$
$\chi^{ai}_{bj} =$	$-r^{ae}_{mj} I^{mi}_{be} + r^e_j I^{ia}_{eb}$
$\chi^{aj}_{ib} =$	$2r^{ae}_{im} I^{mj}_{eb} - r^{ea}_{im} I^{mj}_{eb} - r^{ae}_{im} I^{mj}_{be} + r^e_i I^{ja}_{be}$

a) All elements of $\bar{H}_N$ ($I^{pq\ldots}_{rs\ldots}$) are defined in Table 4.1 whereas χ'^i_j - in Table 4.2.

from the $\hat{R}_3$ operator (all r amplitudes equipped with six indices). In the case of the $(\bar{H}_N R)^{ab}_{ij}$ amplitudes we have of course terms engaging R_3 amplitudes but in addition we have terms with modified $\bar{H}_N$ elements, e.g., I'^{ab}_{ej} is replaced with I^{ab}_{ej}, the latter element includes contribution from $\hat{T}_3$.

Within the CCSDT model the factorization procedure, i.e., replacing the contraction among $\hat{H}_N$ and $\hat{T}_k$ operators with contractions engaging the $\hat{R}$ operators is much more important since it makes it possible to avoid difficult three-body and four-body elements of $\bar{H}_N$.

4.3.3 EE-EOM-CC Results

To illustrate the performance of the EE-EOM-CC method we selected two examples presented in Tables 4.5 and 4.6. In Table 4.5 we collected values of the vertical excitation energies obtained at the EOM-CCSD and EOM-CCSDT levels for the N_2 and CO molecules. The used basis set (aug-cc-pVQZ) is large enough to be sure that the errors due to the basis set limitations are rather small. All computed excited states are dominated by singly excited configurations hence the effect of connected triples is meaningful but not dramatic. Comparing the mean absolute error (MAE) obtained with respect to the experimental values we notice a significant improvement: 0.259 eV reduced to 0.031 eV for the N_2 molecule and for the CO molecule: 0.245 eV reduced to 0.073 eV. The deviations from experiment quoted in parenthesis indicate a very good performance of the EE-EOM-CCSDT scheme even for high-lying excited states. The largest error occurs for the $^1\Sigma^+$ state of the CO molecule.

As a second example we selected results obtained for the ozone molecule. Here the applied basis set is rather poor (POL1) hence the results are not that close to the experiment as in previous cases. We should pay attention to the large error obtained for the first excited singlet state of symmetry A_1 which is dominated by double excitation (configuration obtained by moving an electron from $6a_1$ to $2b_1$ and from $4b_2$ to $2b_1$). At the CCSD level we obtain the excitation energy of 9.95 eV while after inclusion of triples it goes down to 4.93 eV compared to the experimental value of 4.5 eV. Thus the general conclusion is that connected triples must be included to the model if the configurations dominated by the doubles dominate the excited states.

Table 4.5 Vertical excitation energies[a] (in eV) of the N_2 and CO molecules with EOM-CC methods in aug-cc-pVQZ[b] basis set compared to experimental data (deviations from experiment in parenthesis).

	State sym.	EOM-CC SD	EOM-CC SDT	Exp.[c]		State sym.	EOM-CC SD	EOM-CC SDT	Exp.[c]
N_2	$^1\Pi_g$	9.504	9.329	9.34	*CO*	$^1\Pi$	8.656	8.497	8.52
		(0.164)	(−0.011)				(0.136)	(−0.023)	
	$^1\Sigma_u^-$	10.131	9.897	9.88		$^1\Sigma^-$	10.171	9.979	9.90
		(0.251)	(0.017)				(0.271)	(0.079)	
	$^1\Delta_u$	10.568	10.301	10.27		$^1\Delta$	10.306	10.112	10.13
		(0.298)	(0.031)				(0.176)	(−0.017)	
	$^1\Pi_u$	13.371	13.114	13.05		$^1\Sigma^+$	11.178	10.952	10.78
		(0.321)	(0.064)				(0.398)	(0.172)	
	MAE[e]	0.259	0.031				0.245	0.073	

a) From Ref. [66].
b) Valence electrons correlated. Ref. [67].
c) From Ref. [68].
d) Mean absolute error.

Table 4.6 Vertical excitation energies (eV) of the O_3 molecule with EOM-CC methods and POL1[a] basis set (from Ref. [69]).

Nominal state	EOM CCSD	EOM CCSDT	Exp.
2^1A_1 $(6a_1^2 \rightarrow 2b_1^2,\ 4b_2^2 \rightarrow 2b_1^2)$	9.95[b]	4.93[b]	4.5[c]
1^1A_2 $(4b_2 \rightarrow 2b_1)$	2.39[b]	2.06[b]	1.6[d];1.92[e]
1^1B_1 $(6a_1 \rightarrow 2b_1)$	2.40[b]	2.16[b]	2.1[d]
1^1B_2 $(1a_2 \rightarrow 2b_1)$	5.50[b]	5.13[b]	4.86[d]
1^3A_2 $(4b_2 \rightarrow 2b_1)$	1.97[b]	1.81[b]	1.80[d]
1^3B_2 $(1a_2 \rightarrow 2b_1)$	1.48[b]	1.54[b]	1.43; 1.29[d]
1^3B_1 $(6a_1 \rightarrow 2b_1)$	1.77[b]	1.71[b]	1.67[d]

a) Valence electrons correlated. Ref. [70].
b) The optimum (for the given method, i.e., at the CCSD or CCSDT level) geometry has been adopted.
c) Ref. [71].
d) Ref. [72].
e) Ref. [73].

4.4 Ionized States: IP-EOM-CC

The IP realization of EOM-CC theory is introduced by the appropriate definition of the $\hat{R}$ operator. The principal component of the latter is an operator removing the electron from the system, e.g., singly-ionized configurations are obtained by the action of the $\hat{a}_i$ annihilation operator on the $|\Phi_0\rangle$:

$$\hat{a}_i|\Phi_0\rangle = |\Phi_i\rangle$$

i.e., by removing an electron from the ith orbital and with $|\Phi_i\rangle$ representing the configuration containing $N-1$ electrons.

Thus the principal target of the IP-EOM-CC scheme is a description of the electronic structure of cations adopting as a reference system the closed-shell neutral molecule, i.e., performing calculations for the, e.g., H_2O molecule we get results relating to the H_2O^+ cation. However, an inverse situation is also possible. For instance, in order to study the electronic structure of the open-shell neutral radical e.g., CN we may do IP-EOM-CC calculations for the closed-shell anion CN^-. This strategy offers a way to avoid calculations based on a potentially spin-contaminated UHF reference.

4.4.1 IP-EOM-CCSD Model

Within the CCSD model the $\hat{R}$ operator is composed of two terms:

$$\hat{R}(k)_{SD} = \hat{R}_1(k) + \hat{R}_2(k) = \sum_{i}^{N_o} r_i(k)\hat{a}_i + \sum_{a}^{N_v}\sum_{ij}^{N_o} r_{ij}^{a}(k)\hat{a}_a^{\dagger}\hat{a}_j\hat{a}_i$$

where $\hat{R}_1$ is responsible for the ionization and $\hat{R}_2$ – for ionization and single excitation.

In Table 4.7 we list the contributions to the $(\bar{H}_N R)$ target amplitudes. Similarly as in the previous case (EE) most of the terms are the results of contractions between the $\bar{H}_N$ elements and R amplitudes. Two terms engaging three-body $\bar{H}_N$ elements were replaced with their factorized counterparts, see the $F_{ij}{}^{b}$ terms.

4.4.2 IP-EOM-CCSDT Model

Within the CCSDT variant the $\hat{R}$ operator includes the $\hat{R}_3$ component responsible for ionization and double excitation.

$$\hat{R}(k)_{SDT} = \hat{R}(k)_{SD} + \frac{1}{2}\sum_{ab}\sum_{ijl} r_{ijl}^{ab}(k)\hat{a}_a^{\dagger}\hat{a}_b^{\dagger}\hat{a}_l\hat{a}_j\hat{a}_i$$

Table 4.7 The IP-EOM-CCSD equations in the orbital formalism.

		Expression[a)]
$(\bar{H}_N R)_i$	=	$-r_m I_i^m - 2r_{mn}{}^{e}I_{ie}^{mn} + r_{mn}{}^{e}I_{ei}^{mn} + 2r_{im}{}^{e}I_e^m - r_{mi}{}^{e}I_e^m$
$(\bar{H}_N R)_{ij}{}^{b}$	=	$-r_m I_{ij}^{mb} - r_{im}{}^{b}I_j^m - r_{mj}{}^{b}I_i^m + r_{ij}{}^{e}I_e^b + r_{mn}{}^{b}I_{ij}^{mn} + 2r_{im}{}^{e}I_{ej}^{mb} - r_{mi}{}^{e}I_{ej}^{mb} - r_{im}{}^{e}I_{ej}^{bm}$
		$-r_{mj}{}^{e}I_{ie}^{mb} + F_{ij}{}^{b}$
$F_{ij}{}^{b}$ b)	=	$-2r_{mn}{}^{e}I_{iej}^{mnb} + r_{nm}{}^{e}I_{jei}^{bnm}$
$F_{ij}{}^{b}$ c)	=	$t_{ji}^{be}\chi_e$
χ_a	=	$-2r_{no}{}^{e}v_{ae}^{no} + r_{no}{}^{e}v_{ea}^{no}$

a) All elements of $\bar{H}_N$ ($I_{rs\ldots}^{pq\ldots}$) are defined in Table 4.1.
b) Standard version.
c) Factorized version.

Table 4.8 The IP-EOM-CCSDT equations in the orbital formalism (additional terms not included in the IP-EOM-CCSD model).

Expression[a)]		
$(\bar{H}_N R)_i$	=	$2r_{imn}^{\ \ ef}v_{ef}^{mn} - 2r_{min}^{\ \ ef}v_{ef}^{mn} - r_{imn}^{\ \ ef}v_{fe}^{mn} + r_{nmi}^{\ \ fe}v_{fe}^{nm}$
$(\bar{H}_N R)_{ij}^{\ \ b}$	=	$-r_m I_{ij}^{mb} + 2r_{imj}^{\ \ ef}I_{ef}^{mb} - r_{ijm}^{\ \ ef}I_{ef}^{mb} - r_{mij}^{\ \ ef}I_{ef}^{mb} - 2r_{imn}^{\ \ eb}I_{ej}^{mn} + r_{imn}^{\ \ be}I_{ej}^{mn} + r_{min}^{\ \ eb}I_{ej}^{mn}$
		$-2r_{mnj}^{\ \ eb}I_{ie}^{mn} + r_{nmj}^{\ \ eb}I_{ie}^{mn} + r_{mjn}^{\ \ eb}I_{ie}^{mn} + 2r_{imj}^{\ \ eb}I_e^m - r_{ijm}^{\ \ eb}I_e^m - r_{mij}^{\ \ eb}I_e^m + F_{ij}^{\ \ b}$
$(\bar{H}_N R)_{ijk}^{\ \ bc}$	=	$P(jb/kc)[r_{ij}^{\ \ e}I_{ek}^{bc} - r_{im}^{\ \ b}I_{jk}^{mc} - r_{mk}^{\ \ c}I_{ij}^{mb} + r_{mn}^{\ \ b}I_{ijk}^{mnc}] - r_{mjk}^{\ \ bc}I_i^m - P(jb/kc)[r_{imk}^{\ \ bc}I_j^m$
		$-r_{ijk}^{\ \ ec}I_e^b - r_{mnk}^{\ \ bc}I_{ij}^{mn}] + r_{imn}^{\ \ bc}I_{jk}^{mn} + r_{ijk}^{\ \ ef}I_{ef}^{bc} - P(jb/kc)[r_{mjk}^{\ \ ec}I_{ie}^{mb} + r_{ijm}^{\ \ ec}I_{ek}^{bm}$
		$+r_{imk}^{\ \ ec}I_{ej}^{bm} - 2r_{imk}^{\ \ ec}I_{ej}^{mb} + r_{ijm}^{\ \ eb}I_{ek}^{mc} + r_{mik}^{\ \ ec}I_{ej}^{mb}] + F_{ijk}^{\ \ bc}$
$F_{ijk}^{\ \ bc}$ b)	=	$-r_m I_{ijk}^{mbc} + 2r_{im}^{\ \ e}I_{ejk}^{mbc} - r_{mi}^{\ \ e}I_{ejk}^{mbc} - P(jb/kc)[r_{im}^{\ \ e}I_{ejk}^{bmc} + r_{mj}^{\ \ e}I_{iek}^{mbc}] - 2r_{nm}^{\ \ e}I_{iekj}^{nmcb}$
		$+r_{mn}^{\ \ e}I_{eikj}^{mncb} + P(jb/kc)[2r_{imj}^{\ \ ef}I_{efk}^{mbc} - r_{imj}^{\ \ ef}I_{efk}^{bmc} - r_{mij}^{\ \ ef}I_{efk}^{mbc} - 2r_{imn}^{\ \ eb}I_{ejk}^{mnc}$
		$+r_{imn}^{\ \ eb}I_{jek}^{mnc} + r_{min}^{\ \ eb}I_{ejk}^{mnc} - 2r_{mnj}^{\ \ eb}I_{iek}^{mnc} + r_{nmj}^{\ \ eb}I_{iek}^{mnc} + r_{mnj}^{\ \ be}I_{ike}^{mcn}]$
$F_{ijk}^{\ \ bc}$ c)	=	$-r_m I'^{mbc}_{ijk} + P(jb/kc)\, t_{jik}^{bef}\chi_{ef}^{\ \ c} + 2r_{im}^{\ \ e}I'^{mbc}_{ejk} - r_{mi}^{\ \ e}I'^{mbc}_{ejk} - P(jb/kc)[r_{im}^{\ \ e}I'^{bmc}_{ejk}$
		$+r_{mj}^{\ \ e}I'^{mbc}_{iek} + t_{jmi}^{bce}\chi_{ke}^{\ \ m}] - t_{kjm}^{cbe}\chi_{ei}^{\ \ m} + 2t_{jkm}^{bce}\chi_{ei}^{m} - P(jb/kc)\, t_{jmk}^{bce}\chi_{ei}^{m} + t_{ijk}^{ebc}\chi_e$
		$-P(jb/kc)[t_{mk}^{bc}\chi_{ij}^{\ \ m} - t_{jk}^{ec}\chi_{ie}^{\ \ b} - t_{ik}^{ec}\chi_{ej}^{\ \ b}]$
χ_{jk}^{i}	=	$2r_{knj}^{\ \ ef}v_{ef}^{ni} - r_{knj}^{\ \ ef}v_{ef}^{in} - r_{nkj}^{\ \ ef}v_{ef}^{ni}$
$\chi_{ac}^{\ \ b}$	=	$-r_n I_{ac}^{nb}$
χ_{bi}^{a}	=	$-2r_{ino}^{\ \ ea}v_{eb}^{no} + r_{ino}^{\ \ ea}v_{be}^{no} + r_{nio}^{\ \ ea}v_{eb}^{no}$
$\chi_{aj}^{\ \ i}$	=	$-r_{nj}^{\ \ e}v_{ae}^{ni}$
$\chi_{bi}^{\ \ a}$	=	$-2r_{noi}^{\ \ ea}v_{be}^{no} + r_{noi}^{\ \ ea}v_{eb}^{no} + r_{noi}^{\ \ ae}v_{be}^{no}$
$\chi_{ja}^{\ \ i}$	=	$2r_{jn}^{\ \ e}v_{ea}^{ni} - r_{jn}^{\ \ e}v_{ae}^{ni} - r_{nj}^{\ \ e}v_{ae}^{in}$

a) See footnote to Table 4.7.
b) Standard version.
c) Factorized version.

Following the presentation manner introduced in the EE part, in Table 4.8 only those contributions to the $(\bar{H}_N R)_i$ and $(\bar{H}_N R)_{ij}^{\ \ b}$ target amplitudes which originate from the $\hat{R}_3$ operator are shown. Additionally in the $(\bar{H}_N R)_{ij}^{\ \ b}$ equation the incomplete I'^{mb}_{ij} is replaced with its complete counterpart I_{ij}^{mb}. Table 4.8 contains also all terms contributing to the full $(\bar{H}_N R)_{ijk}^{\ \ bc}$ equation.

4.4.3 IP-EOM-CC Results

In Table 4.9 we see computed values for the ionization potentials of three molecules: N_2, CO and F_2. The employed basis set is aug-cc-pVQZ, the assumed bond length are given in table. The results computed at the CCSDT level are in excellent agreement with the available experimental data. For the N_2 molecule the CCSD error of 0.15 eV goes down to less than 0.01 eV upon inclusion of triples; the same is observed for the CO molecule: CCSD is off by 0.25 eV, CCSDT is right on the target. In the worst case, i.e., for the F_2 molecule we observe the reduction from 0.08 eV to 0.04 eV.

In the second example, Table 4.10, we show excitation energies for the cations: N_2^+, CO^+ and F_2^+ computed on the basis of the RHF reference calculated for the neutral molecules. The triples contribution to the excitation energies is also meaningful. For the $^2\Pi_u$ and $^2\Sigma_u$ states of N_2^+ the

Table 4.9 Vertical ionization potentials[a] (eV) with IP-EOM-CC methods (N_2 : R=1.097685 Å, CO: R=1.128323 Å, F_2 : R=1.41193 Å).

		aug-cc-pVQZ[b]		
Molecule	**Sym.**	**EOM-CCSD**	**EOM-CCSDT**	**Exp.[c]**
N_2	σ_g	15.75	15.60	15.60
CO	σ	14.26	14.01	14.01
F_2	π_g	15.75	15.79	15.83

a) From Refs. [26].
b) Ref. [67].
c) Refs. [74].

Table 4.10 Computed vertical excitation energies (eV) with IP-EOM-CC methods.

		aug-cc-pVQZ		
Cation	**Sym.**	**EOM-CCSD**	**EOM-CCSDT**	**Exp.[a]**
N_2^+	$^2\Pi_u$	1.59	1.45	1.38
	$^2\Sigma_u^-$	3.23	3.19	3.18
CO^+	$^2\Pi$	2.92	3.08	2.9
	$^2\Sigma^+$	5.63	5.65	5.71
F_2^+	$^2\Pi_u$	3.34	3.18	2.97
	$^2\Sigma_g^+$	5.52	5.39	5.27

a) Refs. [74].

deviations of the excitation energies from experiment are reduced from 0.21 eV to 0.07 eV and from 0.05 eV to 0.01 eV for both states, respectively. The excitation energies of the two states of CO^+ shown in Table 4.10, $^2\Pi$ and $^2\Sigma$, are already quite well reproduced at the CCSD level (deviations of 0.02 and 0.08 eV, respectively). The full CCSDT in the first case increases the error to 0.18 eV and in the second case reduces to 0.06 eV. In the F_2^+ case the CCSD values are off by 0.37 eV and 0.25 eV for the two states shown in Table 4.10 and improved due to inclusion of triples by 0.16 and 0.13 eV, respectively.

4.5 Electron-Attached States: EA-EOM-CC

The EA-EOM-CC realization of the EOM-CC method is fully analogous in its canonical equations to the IP counterpart. However, in the practical realization the factorization procedure is somewhat different hence the working equations are modified. The single-electron attached configurations are obtained by creating an electron in one of the virtual orbitals with the operator $\hat{a}_a^\dagger$:

$$\hat{a}_a^\dagger|\Phi_0\rangle = |\Phi^a\rangle$$

where $|\Phi^a\rangle$ is a configuration with $N+1$ electrons.

4.5.1 EA-EOM-CCSD Model

Within the CCSD model the $\hat{R}$ operator is composed of two terms:

$$\hat{R}(k)_{SD} = \sum_{a} r^{a}(k)\hat{a}_{a}^{\dagger} + \sum_{ab}\sum_{i} r_{i}^{ab}(k)\hat{a}_{a}^{\dagger}\hat{a}_{b}^{\dagger}\hat{a}_{i}$$

The last term represents the operator placing the additional electron in the virtual level a with an accompanying single excitation.

Algebraic contributions to the $(\bar{H}_N R)^a$ and $(\bar{H}_N R)_i^{ab}$ target amplitudes are shown in Table 4.11. Similarly as in the IP case the factorized contributions are represented by the $F_i^{ab\ c)}$ term in order to avoid employing the three-body $\bar{H}_N$ elements.

4.5.2 EA-EOM-CCSDT Model

Within the CCSDT model the $\hat{R}$ operator includes an additional term responsible for attaching an electron and accompanying double excitation:

$$\hat{R}(k)_{SDT} = \hat{R}(k)_{SD} + \frac{1}{2}\sum_{abc}\sum_{ij} r_{ij}^{abc}(k)\hat{a}_{a}^{\dagger}\hat{a}_{b}^{\dagger}\hat{a}_{c}^{\dagger}\hat{a}_{j}\hat{a}_{i}$$

Algebraic contributions to the target $(\bar{H}_N R)^a$, $(\bar{H}_N R)_i^{ab}$ amplitudes due to the $\hat{T}_3$ and $\hat{R}_3$ operators are presented in Table 4.12 together with the full set of terms occurring in the $(\bar{H}_N R)_{ij}^{abc}$ equation. However the EA variant is computationally much more demanding than its IP equivalent due to the fact that the number of EA amplitudes is of the order of $N_o^2 N_v^5$ compared to the $N_o^5 N_v^2$ and usually the number of virtual levels is much larger than the number of occupied ones. Hence the factorization introduced in the EA-EOM-CCSDT approach is much deeper than in the IP scheme.

4.5.3 EA-EOM-CC Results

To illustrate the usefulness of the EA-EOM-CC approach in the calculations of the excited states we show computed values of the atomic energy levels for the open-shell atoms Li and Na in Table 4.13. The RHF-based reference calculations were done for the closed-shell ions Li^+ and Na^+. Applying the EA-EOM-CC scheme we get the energies of the neutral open-shell atoms. In the case of Li both the CCSD and CCSDT results are very close to experiment with errors oscillating around 0.001 eV. Note that the EA-EOM-CCSDT results for a Li atom are identical to the FCI values. The triples effect is small, of the order of 0.0001 eV, but in all cases goes in the right direction. For the Na atom

Table 4.11 The EA-EOM-CCSD equations in the orbital formalism.

Expression[a)]		
$(\bar{H}_N R)^a$	=	$r^e I_e^a + 2r_m^{ef} I_{ef}^{ma} - r_m^{ef} I_{ef}^{am} + 2r_m^{ea} I_e^m - r_m^{ea} I_e^m$
$(\bar{H}_N R)_i^{ab}$	=	$r^e I_{ie}^{\prime ab} + r_i^{eb} I_e^a + r_i^{ae} I_e^b - r_m^{ab} I_i^m + r_i^{ef} I_{ef}^{ab} + 2r_m^{eb} I_{ie}^{am} - r_m^{be} I_{ie}^{am} - r_m^{eb} I_{ei}^{am} - r_m^{ae} I_{ie}^{mb} + F_i^{ab}$
$F_i^{ab\,b)}$	=	$2r_m^{ef} I_{ief}^{amb} - r_m^{ef} I_{efi}^{bma}$
$F_i^{ab\,c)}$	=	$-t_{im}^{ab} \chi'^m$
χ'^i	=	$2r_m^{fg} v_{fg}^{mi} - r_m^{fg} v_{fg}^{im}$

a) See footnote to Table 4.7.
b) Standard version.
c) Factorized version.

Table 4.12 The EA-EOM-CCSDT equations in the orbital formalism.

Expression[a]		
$(\bar{H}_N R)^a$	=	$+2r^{efa}_{mn}v^{mn}_{ef} - 2r^{eaf}_{mn}v^{mn}_{ef} - r^{efa}_{mn}v^{mn}_{fe} + r^{aef}_{mn}v^{mn}_{ef}$
$(\bar{H}_N R)^{ab}_i$	=	$r^e I^{ab}_{ie} + r^{eb}_i I^a_e + 2r^{efb}_{im} I^{am}_{ef} - r^{efb}_{mi} I^{am}_{ef} - r^{ebf}_{im} I^{am}_{ef} + 2r^{aef}_{im} I^{mb}_{ef} - r^{aef}_{im} I^{bm}_{ef} - r^{eaf}_{im} I^{mb}_{ef}$
		$-2r^{aeb}_{mn} I^{mn}_{ie} + r^{eab}_{mn} I^{mn}_{ie} + r^{abe}_{mn} I^{mn}_{ie} + 2r^{aeb}_{im} I^m_e - r^{eab}_{im} I^m_e - r^{abe}_{im} I^m_e$
$(\bar{H}_N R)^{abc}_{ij}$	=	$P(ia/jb)[r^{ec}_j I^{ab}_{ie} + r^{be}_j I^{ac}_{ie} - r^{bc}_m I^{am}_{ij}] + r^{abe}_{ij} I^c_e + P(ia/jb)[r^{aec}_{ij} I^b_e - r^{abc}_{im} I^m_j$
		$+r^{aef}_{ij} I^{bc}_{ef}] + r^{efc}_{ij} I^{ab}_{ef} + r^{abc}_{mn} I^{mn}_{ij} - P(ia/jb)[r^{abe}_{im} I^{mc}_{je} + r^{aec}_{mj} I^{mb}_{ie} + r^{aec}_{im} I^{mb}_{je}$
		$-2r^{aec}_{im} I^{bm}_{je} + r^{ebc}_{jm} I^{am}_{ie} + r^{ace}_{im} I^{bm}_{je}] + F^{abc}_{ij}$
$F^{abc\ b)}_{ij}$	=	$r^e I^{abc}_{ije} + 2r^{ec}_m I^{abm}_{ije} - r^{ce}_m I^{abm}_{ije} - P(ia/jb)[r^{ec}_m I^{abm}_{iej} + r^{be}_m I^{amc}_{ije} - r^{ef}_j I^{abc}_{ief}]$
		$+2r^{ef}_m I^{abmc}_{ijef} - r^{ef}_m I^{abcm}_{ijef} + P(ia/jb)[2r^{efc}_{jm} I^{abm}_{ief} - r^{efc}_{mj} I^{abm}_{ief} - r^{ecf}_{jm} I^{abm}_{ief}$
		$+2r^{aef}_{im} I^{bmc}_{jef} - r^{aef}_{im} I^{bcm}_{jef} - r^{bef}_{mj} I^{amc}_{ief} - 2r^{bec}_{mn} I^{amn}_{ije} + r^{ebc}_{mn} I^{amn}_{ije} + r^{bce}_{mn} I^{amn}_{ije}]$
$F^{abc\ c)}_{ij}$	=	$-t^{abc}_{ijm}\chi^m - P(ia/jb)\,[t^{ab}_{im}\chi^{mc}_j + t^{ac}_{im}\chi^{bm}_j - t^{ae}_{ij}\chi^{bc}_e + t^{aec}_{ijm}\chi^{bm}_e] - t^{abe}_{ijm}\chi^{cm}_e$
		$+2t^{abe}_{ijm}\chi^{mc}_e - P(ia/jb)[t^{abe}_{imj}\chi^{mc}_e - t^{abc}_{imn}\chi^{mn}_j]$
χ^i	=	$\chi'i + r^f I^i_f$
χ^{ia}_j	=	$r^f I^{ia}_{jf} + r^{fg}_j I^{ia}_{fg} - r^{fa}_m I^{mi}_{jf} + 2r^{fa}_m I^{im}_{jf} - r^{af}_m I^{im}_{jf} + 2r^{fga}_{jm} v^{im}_{fg} - r^{fga}_{jm} v^{mi}_{fg} - r^{fag}_{jm} v^{im}_{fg}$
χ^{aj}_i	=	$r^f I^{aj}_{if} + r^{fg}_i I^{aj}_{fg} - r^{af}_m I^{mj}_{if} + 2r^{afg}_{im} v^{mj}_{fg} - r^{afg}_{im} v^{jm}_{fg} - r^{fag}_{im} v^{mj}_{fg}$
χ^{ab}_c	=	$r^f I^{ab}_{cf} - r^{af}_m I^{mb}_{cf} - r^{fb}_m I^{am}_{fc} + 2r^{fb}_m I^{am}_{cf} - r^{bf}_m I^{am}_{cf} - 2r^{afb}_{mn} v^{mn}_{cf} + r^{afb}_{mn} v^{mn}_{fc} + r^{abf}_{mn} v^{mn}_{cf}$
χ^{ai}_b	=	$r^f I^{ai}_{bf} - r^{af}_m v^{mi}_{bf}$
χ^{ib}_a	=	$r^f I^{ib}_{af} - r^{fb}_m v^{mi}_{af} + 2r^{fb}_m v^{im}_{af} - r^{bf}_m v^{im}_{af}$
χ^{ik}_j	=	$r^f I^{ik}_{jf} + r^{fg}_j v^{ik}_{fg}$

a) See footnote to Table 4.7. For definition of intermediate (χ'^i) see Table 4.11.
b) Standard version.
c) Factorized version.

Table 4.13 Atomic energy levels (eV) within EA-EOM-CC methods and the unANO-RCC+ basis set (from Ref. [75]).

Atom	Sym.	EA-EOM-CCSD	Δ[a]SD	EA-EOM-CCSDT	Δ[a]SDT	Exp.[c]
Li	^2P	1.8492	0.0014	1.8491[b]	0.0013	1.8478
	^2S	3.3707	−0.0024	3.3709[b]	−0.0022	3.3731
	^2P	3.8339	−0.0004	3.8341[b]	−0.0002	3.8343
	^2D	3.8757	−0.0029	3.8761[b]	−0.0025	3.8786
Na	^2P	2.0871	−0.0152	2.0963	−0.0060	2.1023
	^2S	3.1721	−0.0193	3.1826	−0.0088	3.1914
	^2D	3.5939	−0.0231	3.6067	−0.0103	3.6170
	^2P	3.7311	−0.0215	3.7431	−0.0095	3.7526

a) Deviation from experiment.
b) Equivalent to FCI value.
c) Refs. [76, 77].

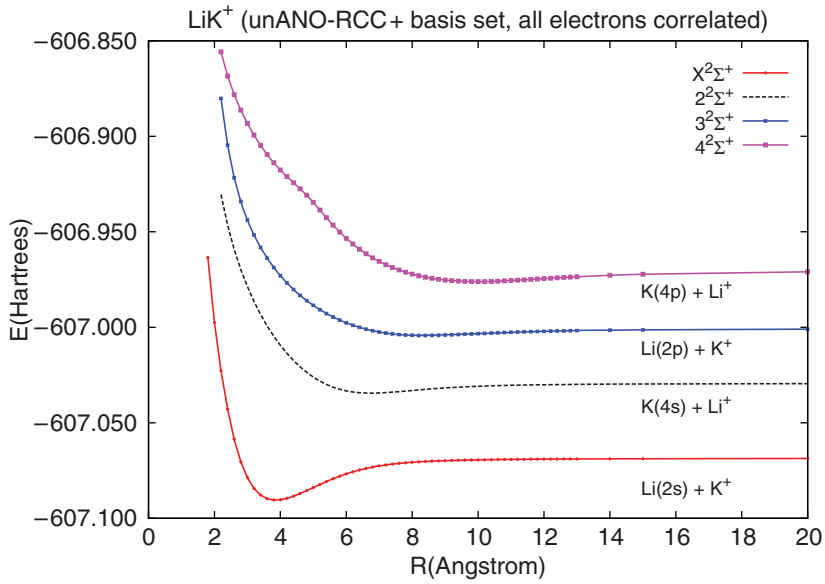


Figure 4.1 Potential energy curves of the LiK^+ ion with the EA-EOM-CCSD method in unANO-RCC+ basis set (from Ref. [75]).

the deviations from experiment are larger, see respective columns (ΔSD, ΔSDT) and also the triples effect is larger amounting in all considered cases to ca. 0.01 eV.

As a second example demonstrating the possible use of the EA-EOM-CC scheme in the characterization of excited states we selected calculations of the potential energy curves (PEC) for the LiK^+ ion. The adopted computational strategy is as follows: we want to compute PEC for the open-shell system by a method based on the RHF function. As the reference system we adopted the double positive cation LiK^{2+} which dissociates $LiK^{2+} \rightarrow Li^+ + K^+$ into the closed-shell fragments Li^+ and K^+ in the whole region of interatomic distances. By doing EA-EOM calculations for each point of the PEC we can plot the energy values for the LiK^+ ion. The adopted strategy ensures a generation of the smooth energy curves as shown in Figure 4.1.

4.6 Doubly-Ionized States: DIP-EOM-CC

This variant of EOM-CC is aimed at the evaluation of double ionization potentials and the energies of the excited electronic states of doubly ionized system. Doubly ionized states can be accessed experimentally, hence the performance of the DIP-EOM-CC method can be easily verified. Before we introduce the respective equations we would like to indicate another possibility to use DIP-EOM-CC results. For instance, adopting the system A with two attached electrons, i.e., A^{2-} as a reference, and performing the DIP-EOM-CC calculations we obtain results pertaining to the neutral molecule A. This strategy could be of particular interest in situations when the ground state (in this case it is more appropriate to use the reference state) CC calculations are easier to obtain for A^{2-} than for A. Consider, e.g., the dissociation of the F_2 molecule. The proper description of the potential energy curve encounters some difficulties since the closed-shell molecule dissociates

(we consider a homolytic process) into open-shell fragments:

$$F_2 \rightarrow F + F$$

To account properly for the open-shell character of the products we need to use the UHF approach with its well known deficiencies. The alternative route would be to adopt the F_2^{2-} ion dissociating into the closed-shell fragments as a reference system:

$$F_2^{2-} \rightarrow F^- + F^-$$

(F^- being isoelectronic with the neon atom). When computing PEC with the DIP-EOM-CC scheme for F_2^{2-} system we can use the RHF based approach for the whole range of interatomic distances computing at each point the energy of the neutral F_2 molecule. Below we will recall the other situations for which the DIP-EOM-CC calculations seem particularly suitable. As before we will discuss two CC models: CCSD and CCSDT.

4.6.1 DIP-EOM-CCSD Model

Within the CCSD model the $\hat{R}$ operator is represented by one term only, responsible for the removal of two electrons from the system:

$$\hat{R}(k)_{SD} = \frac{1}{2}\sum_{ij} r_{ij}(k)\hat{a}_j\hat{a}_i$$

Consequently the resulting equation for the $(\bar{H}_N R)_{ij}$ target amplitude is rather straightforward, see Table 4.14.

4.6.2 DIP-EOM-CCSDT Model

Within the CCSDT model $\hat{R}(k)_{SDT}$ is defined as:

$$\hat{R}(k)_{SDT} = \hat{R}_{SD}(k) + \frac{1}{2}\sum_{ijl}\sum_{a} r_{ijl}{}^{a}(k)\hat{a}_a^\dagger\hat{a}_l\hat{a}_j\hat{a}_i$$

The $\hat{R}(k)_{SDT}$ given above includes the component responsible for the annihilation of two electrons connected with a single excitation.

The resulting equations are more complex and in the standard form engage the three-body elements of $\bar{H}_N$ eliminated in the factorized form and replaced with the χ intermediates, see Table 4.15.

Moreover in the case of DIP and DEA the hybrid method (acronyms DIP-EOM-CCSD', DEA-EOM-CCSD') are considered. In these variants the ground state is solved at the CCSD level (scaling n^6) instead of CCSDT (scaling n^8) and in EOM part (scaling n^6) the R_2 and R_3 equations are taken as in the full CCSDT method (see Refs. [33, 36] for more details). These new variants give results of comparable quality to the full DIP-EOM-CCSDT and DEA-EOM-CCSDT methods but with lower scaling.

Table 4.14 The DIP-EOM-CCSD equation in the orbital formalism.

Expression[a)]
$(\bar{H}_N R)_{ij} = \quad P(i/j)[-r_{im}I_j^m + \frac{1}{2}r_{mn}I_{ij}^{mn}]$

a) See footnote to Table 4.7.

Table 4.15 The DIP-EOM-CCSDT equations in the orbital formalism. (additional terms not included in the DIP-EOM-CCSD model).

Expression[a]	
$(\bar{H}_N R)_{ij} =$	$P(i/j)[r_{mn} I^{mn}_{ij} - 2r_{imn}\ {}^{e}I^{mn}_{je} + r_{imn}\ {}^{e}I^{mn}_{ej} + r_{mnj}\ {}^{e}I^{mn}_{ie} + r_{ijm}\ {}^{e}I^{m}_{e} - r_{imj}\ {}^{e}I^{m}_{e}]$
$(\bar{H}_N R)_{ijk}{}^{a} =$	$P(i/j)[-r_{im} I^{ma}_{jk} + \frac{1}{2} r_{ijk}\ {}^{e}I^{a}_{e} - \frac{1}{2} r_{ijm}\ {}^{a}I^{m}_{k} - r_{mjk}\ {}^{a}I^{m}_{i} + \frac{1}{2} r_{mnk}\ {}^{a}I^{mn}_{ij}$
	$+ r_{imn}\ {}^{a}I^{mn}_{jk} - r_{imk}\ {}^{e}I^{ma}_{je} - \frac{1}{2} r_{ijm}\ {}^{e}I^{ma}_{ke} + r_{ijm}\ {}^{e}I^{ma}_{ek} - r_{imj}\ {}^{e}I^{ma}_{ek}] + F_{ijk}{}^{a}$
$F_{ijk}{}^{ab)} =$	$P(i/j)[\frac{1}{2} r_{mn} I^{mna}_{ijk} - 2r_{imn}\ {}^{e}I^{mna}_{jek} + r_{imn}\ {}^{e}I^{mna}_{ejk} + r_{mnj}\ {}^{e}I^{mna}_{iek}]$
$F_{ijk}{}^{ac)} =$	$P(i/j)[\frac{1}{2} t^{efa}_{ijk} \chi_{ef} + t^{ea}_{jk} \chi_{ie}]$
$\chi_{ab} =$	$r_{mn} v^{mn}_{ab}$
$\chi_{ia} =$	$r_{mn} I^{mn}_{ia} - r_{imn}\ {}^{e}v^{mn}_{ae} + r_{mni}\ {}^{e}v^{mn}_{ae} + r_{imn}\ {}^{e}v^{mn}_{ea}$

a) See footnote to Table 4.7.
b) Standard version.
c) Factorized version.

4.6.3 DIP-EOM-CC Results

In Table 4.16 we listed vertical double ionization potentials for a set of closed-shell molecules. Note that the results collected here represent the energy difference between the ground singlet state of the neutral molecules and the ground and excited states of their double positive ions. Thus the

Table 4.16 Vertical double ionization potentials[a] (eV) with the EOM-CC methods in the POL1 basis set (from Ref. [78]).

		DIP-EOM			
		CCSD	CCSDT	CCSDT'	
		GS : n^6	*GS* : n^8	*GS* : n^6	
Mol.	Sym.	*EOM* : n^4	*EOM* : n^6	*EOM* : n^6	Exp.[b]
H_2O	$^1A_1$	50.39	41.73	41.63	41.3
	$^1B_1$	51.78	43.24	43.14	42.0
CO	$^1\Sigma^+$	46.72	41.91	41.76	41.7
	$^1\Pi$	47.02	42.71	42.53	42.2
	$^1\Sigma^+$	51.75	45.94	45.72	45.8
C_2H_2	$^1\Delta_g$	38.13	33.88	33.63	33.0
	$^1\Pi_u$	43.71	39.15	38.93	37.6
	$^1\Pi_g$	45.82	41.30	41.09	39.3
C_2H_4	^1A_g	35.59	31.20	30.96	30.1
	^1A_g	37.06	32.79	32.58	32.2
	$^1B_{3u}$	40.66	35.57	35.34	34.0
MAE		6.32	0.93	0.75	

a) Assumed experimental geometry for the ground state.
b) Refs. [79–84].

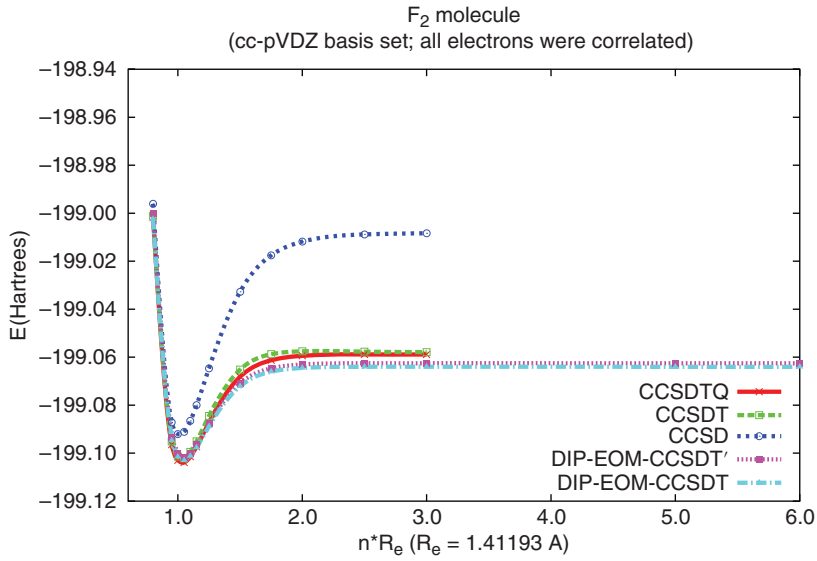


Figure 4.2 Potential energy curves of the ground state of the F_2 molecule within CC and DIP-EOM-CC methods (from Ref. [78]).

energy values listed in the first row at each molecule represent the double ionization potential and can be used to get the ground state energy of the M^{2+} ions. The energy values corresponding to higher states can be used to evaluate the excitation energy for the M^{2+} ions. An important issue is also molecular geometry at which the energy values are computed. If the vertical excitation energy of the doubly ionized M^{2+} is required the DIP calculations must be done for the optimum geometry of the lowest lying state. Comparing the results given in Table 4.16 with available experimental data we may assess the importance of the inclusion of the triples in the DIP calculations. In all cases the triples are critical in recovering a satisfactory agreement with experiment reducing the error from several eV to a few tenths of an eV. This can be seen by comparing the mean absolute error of 6.32 eV for the CCSD variant and 0.73 for the CCSDT one.

Another example of using the the DIP-EOM scheme is shown in Figure 4.2. The plot represents the potential energy curves of the F_2 molecule obtained with the assumption of the DIP strategy discussed above. The reference system is represented by the F_2^{2-} dianion which dissociates into the closed-shell F^- ions isoelectronic with the neon atom. As a result, smooth curves are computed for bond distances ranging from equilibrium to infinity.

4.7 Doubly Electron-Attached States: DEA-EOM-CC

At first sight an evaluation of the states with two attached electrons seems to be of little interest since such states are rarely available experimentally. Hence the direct calculation of the transition energy connected with attaching of two electrons to a neutral molecule has little sense. In the application described below we focus on using the DEA-EOM approach in the inverse manner: i.e., we will apply the DEA scheme to characterize neutral atoms or molecules by adopting the doubly ionized system as the reference. For example, using the standard coupled-cluster or configuration interaction schemes to get the electronic structure of the (open-shell) carbon atom would result

in a broken symmetry solution (unless we use FCC or FCI schemes). In the DEA calculations we assume the doubly ionized cation C^{2+} as the reference system and using DEA-EOM scheme we recover the correct symmetry and energy both of ground and excited states of the neutral carbon atom. Similarly – in analogy to the DIP calculations discussed in the previous section – we may reproduce correct potential energy curves for the Na_2 dimer assuming the closed-shell Na_2^{2+} ion as the reference system dissociating into two closed-shell fragments:

$$Na_2^{2+} \rightarrow Na^+ + Na^+$$

(Na^+ being isoelectronic with the neon atom).

Bearing this in mind it becomes worthwhile to develop the DEA-EOM-CC approach.

4.7.1 DEA-EOM-CCSD Model

Computationally the DEA variant based on the $\hat{R}_2$ operator ($\hat{R}_1$ does not occur both in DIP and DEA variants) is rather straightforward:

$$\hat{R}(k)_{SD} = \frac{1}{2}\sum_{ab} r^{ab}(k)\hat{a}_a^\dagger \hat{a}_b^\dagger$$

The $\hat{R}(k)$ operator places additional electrons on virtual levels a and b with a and b running over all virtuals. In analogy to the DIP variant the $(\bar{H}_N R)^{ab}$ target amplitudes are composed of two terms, see Table 4.17. The size of the problem is significantly larger than in the DIP case.

4.7.2 DEA-EOM-CCSDT Model

We formally introduced the DIP and DEA EOM variants at the CCSD level although they are of little practical use. The usefulness of this method can be shown only at the CCSDT level where the $\hat{R}_3(k)$ operator is added to the $\hat{R}_{SD}(k)$ one:

$$\hat{R}(k)_{SDT} = \hat{R}_{SD}(k) + \frac{1}{2}\sum_{abc}\sum_{i} r^{abc}{}_i(k)\hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_c^\dagger \hat{a}_i$$

Applying a generalized Davidson scheme the target $(\bar{H}_N R)^{abc}{}_i$ amplitudes include the terms given in Table 4.18.

In this case, i.e., at the CCSDT level, the factorization of the DEA contribution is necessary to avoid construction of the three-body $\bar{H}_N$ elements engaging i.a. five virtual levels.

4.7.3 DEA-EOM-CC Results

An example illustrating the usefulness of the DEA-EOM scheme is shown in Table 4.19 containing the term energy values of the carbon atom. We adopt the closed-shell C^{2+} ion ($1s^2 2s^2$ configuration)

Table 4.17 The DEA-EOM-CCSD equation in the orbital formalism.

Expression[a)]	
$(\bar{H}_N R)^{ab} =$	$P(a/b)[r^{ae} I_e^b + \frac{1}{2} r^{ef} I_{ef}^{ab}]$

a) See footnote to Table 4.7.

Table 4.18 The DEA-EOM-CCSDT equations in the orbital formalism. (additional terms not included in the DEA-EOM-CCSD model).

Expression[a]	
$(\bar{H}_N R)^{ab} =$	$P(a/b)[2r^{aef}{}_m I^{bm}_{ef} - r^{aef}{}_m I^{mb}_{ef} - r^{efb}{}_m I^{am}_{ef} + r^{abe}{}_m I^m_e - r^{aeb}{}_m I^m_e]$
$(\bar{H}_N R)^{abc}{}_i =$	$P(a/b)[r^{ae} I^{bc}_{ei} + \frac{1}{2} r^{abe}{}_i I^c_e - \frac{1}{2} r^{abc}{}_m I^m_i + r^{ebc}{}_i I^a_e + \frac{1}{2}\sum_{ef} r^{efc}{}_i I^{ab}_{ef}$
	$+ r^{aef}{}_i I^{bc}_{ef} - r^{aec}{}_m I^{bm}_{ei} - \frac{1}{2} r^{abe}{}_m I^{mc}_{ie} + r^{abe}{}_m I^{mc}_{ei} - r^{aeb}{}_m I^{mc}_{ei}] + F^{abc}{}_i$
$F^{abc}{}_i$ [b] $=$	$P(a/b)[\frac{1}{2} r^{ef} I^{abc}_{efi} + 2r^{aef}{}_m I^{bmc}_{efi} - r^{aef}{}_m I^{mbc}_{efi} - r^{efb}{}_m I^{amc}_{efi}]$
$F^{abc}{}_i$ [c] $=$	$P(a/b)[\frac{1}{2} t^{abc}_{mni} \chi^{mn} - t^{bc}_{mi} \chi^{am}]$
$\chi^{ij} =$	$r^{ef} v^{ij}_{ef}$
$\chi^{ai} =$	$r^{ef} I^{ai}_{ef} + 2r^{aef}{}_m v^{im}_{ef} - r^{efa}{}_m v^{im}_{ef} - r^{aef}{}_m v^{mi}_{ef}$

a) See footnote to Table 4.7.
b) Standard version.
c) Factorized version.

Table 4.19 Term energies (eV) for the C atom using the DEA-EOM-CC methods in the POL1 basis set (from Ref. [85]).

	DEA-EOM			
	CCSD	**CCSDT**	**CCSDT'**	
	***GS* : n^6**	***GS* : n^8**	***GS* : n^6**	
Sym.	***EOM*:n^4**	***EOM* : n^6**	***EOM* : n^6**	**Exp.**[a]
$2p^2\ ^1D$	1.308	1.480	1.481	1.26
$2p^2\ ^1S$	1.771	2.802	2.805	2.68
$2p3s\ ^3P^0$	5.834	7.318	7.320	7.48
$2p3s\ ^1P^0$	5.963	7.520	7.521	7.68
$2p3p\ ^1P$	7.107	8.503	8.504	8.54
$2p3p\ ^3D$	7.228	8.670	8.671	8.64
$2p3p\ ^3S$	7.411	8.869	8.870	8.77
$2p3p\ ^3P$	7.748	9.026	9.027	8.85
$2p3p\ ^1D$	8.051	9.382	9.384	9.00
MAE[b]	1.175	0.154	0.155	

a) Ref. [86].
b) MAE – mean absolute error.

as reference and then by using DEA-EOM schemes we recover the energy of the neutral C atom. The basis set is rather modest hence the results for the EOM-CCSDT scheme are off the experiment by 0.1–0.2 eV. The average error for nine electronic states is 0.154 eV, i.e., by more than 1 eV lower than those for the CCSD method.

In Figure 4.3 we present the potential energy curves for the Na_2 dimer. The adopted reference system is Na_2^{2+} dissociating according to the scheme $Na_2^{2+} \rightarrow Na^+ + Na^+$, i.e., into two closed-shell fragments. The DEA-EOM calculations recover the original process engaging the neutral molecule

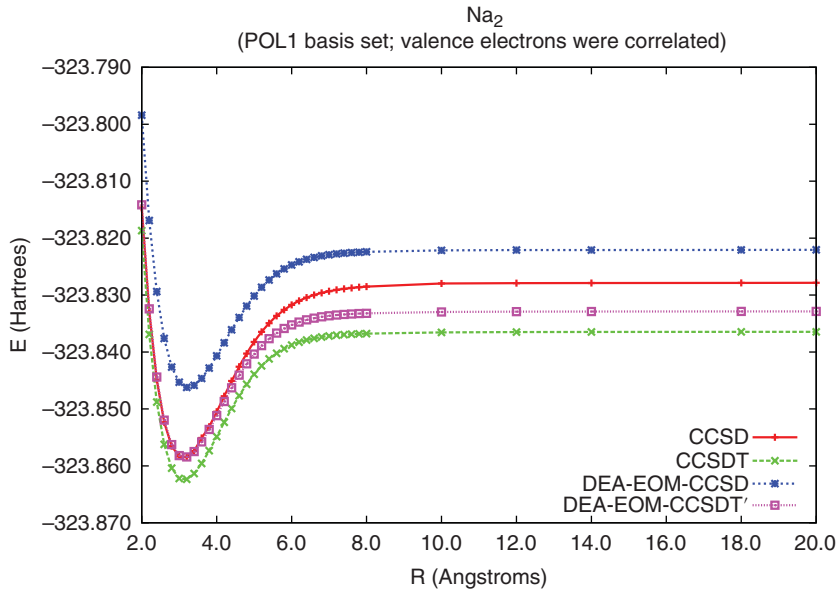


Figure 4.3 Potential energy curves for the ground state of the Na_2 molecule within CC and DEA-EOM-CC methods (from Ref. [85]).

and atoms, i.e., $Na_2 \rightarrow Na + Na$. The obtained PECs correctly reproduce the experimental curves [85].

4.8 Size-Extensivity Issue in the EOM-CC Theory

An attractive feature of the coupled-cluster theory is its size-extensivity property. This means that the total energy of the collection of non-interacting units is a sum of the energies computed separately for each unit. Within EOM-CC theory this problem is more complicated. Let us consider two non-interacting molecules A and B and use the EOM-CC method to get excitation energies (we assume that we have truncated expansion of the $\hat{R}$ operator). If the method is size extensive then all possible types of excitations comply to the condition that we get the same energy for the complex as for the separate calculations for the components. First we want to investigate whether the excitation energies computed separately for the molecules A and B are the same as the excitation energies computed for the AB complex. The EOM-CC provides a positive answer to that problem. If we get from the calculations for the AB complex the energies of the local excitations (i.e., occurring within A or within B unit) then we get the identical results as for the monomer calculations. This feature of the method is called size-intensivity [38, 39, 87–89] and the EOM-CC method is size-intensive.

However, there is a different situation for the charge-transfer (CT) calculations, i.e., when the electron removed from the monomer, say, A (donor), is placed on the virtual level of monomer B (acceptor). For such a CT state the energy at large internuclear separation should be expressed as: $EE = IP(D) + EA(A) - e^2/R_{DA}$, where the IP and EA values are taken from the IP/EA-EOM-CC calculations and the last term represents the electrostatic interaction between the donor and acceptor. For the EOM-CC the expression for CT at the separable limit does not hold. This is rather a small effect but it is a formal deficiency of the theory that should be fixed.

Table 4.20 Behavior of charge transfer excitation in a Be· · · C_2 complex[a] (from Ref. [90]).

		$EE-(IP+EA)$ (in eV)		
		EOM-CC		
R (Å)	$-\frac{e^2}{R}$ (eV)	SD	SDT	SDx
5	−2.880	−2.937	−2.872	−3.012
		(−0.057)	(0.008)	(−0.132)
10	−1.440	−1.375	−1.437	−1.440
		(0.065)	(0.003)	(0.000)
100	−0.1440	−0.0799	−0.1411	−0.1440
		(0.064)	(0.003)	(0.000)
1000	−0.0144	0.0496	−0.0115	−0.0144
		(0.064)	(0.003)	(0.000)

a) In parenthesis: $\Delta = EE-(IP+EA)+\frac{e^2}{R}$ (eV) for large R represents size-extensivity error.

The problem of the behavior of EEs when a system is separated into two non-interacting subsystems has been discussed on the example of the $Be...C_2$ complex (see Ref. [90] for computational details). The results obtained there are presented in Table 4.20. In the columns with headers *SD* and *SDT* we have the difference between the CT excitation energy and the sum of ionization potential and electron affinity ($EE-(IP+EA)$). This value for larger distances should be equal to the electrostatic interaction energy between donor and acceptor (listed in the second column). As we see the size-extensivity error is not large, at the CCSD level it is equal to 0.064 eV while upon inclusion of connected triples it is reduced to 0.003 eV. Thus for higher EOM-CC models the size-extensivity errors are negligible.

In the same table in the last column we have shown that the method denoted as EOM-CCSDx obtained by small modifications of the regular EOM-CCSD scheme provides size-extensive results. The general prescription for such modification relies on the addition of selected terms to the $\bar{H}_N$ elements to eliminate the terms causing the size-extensivity errors (in the diagrammatic language they are called unlinked terms). An exhaustive discussion of the charge-transfer separability (CTS) problem and why it is not satisfied by ordinary EOM-CC or CCLR has been given by Stanton [89].

It should be mentioned here that the method with correct size-extensivity has been proposed by Nooijen and coworkers [91, 92] who considered a modification of the standard EOM-CC approach by introducing a second similarity transformation to the EOM equations (STEOM - Similarity Transformed EOM). It also correctly describes charge-transfer separability. The true remedy for the size-extensivity problem of the EOM-CC approach is to use the multi-reference Fock space CC approach (MR-FS-CC) [7, 9] which by definition provides size-extensive results and correctly describes CT separability since it provides a valence universal wave operator built on an exponential ansatz. Moreover, IH-MR-FS-CCSD (IH - Intermediate Hamiltonian) reduces to STEOM-CCSD when all orbitals are active [90, 93]. However, EOM-CCx is different from both.

The proper formulation of the MR-FS-CC method via the intermediate Hamiltonian formulation in the two valence sectors (i.e., for EE - (1,1), DIP - (0,2) and DEA - (2,0)) offers an easy way to replace the iterative solution of the Bloch equation by direct diagonalization of the properly constructed matrix ("dressed" $\bar{H}_N$ called IH) [66, 78, 85, 93, 94] with the help of the EOM-CC strategy. Moreover,

at the (0,1) (IP) and (1,0) (EA) level the IH formulation reduces the MR-FS-CC methods to the IP-EOM-CC and EA-EOM-CC ones. As is known these vectors give the same eigenvalues with the only difference being that the cluster amplitudes for the (0,1) and (1,0) sectors must be determined from the sets of the IP- and EA-EOM-CC eigenvectors. Thus, the EOM-CC solution can be obtained either by direct diagonalization of the $\bar{H}_N$ matrix within the appropriate configurational subspace or by solving the Bloch equation iteratively (see Ref. [90] for more details).

4.9 Final Remarks

The various computational schemes hidden under the name EOM-CC create a path through which coupled-cluster theory unifies the studies of excited, ionized and electron-attached molecular systems. EOM-CC theory offers a plethora of computational schemes depending on the definition of the $\hat{R}$ operator. All of them have similar structure, i.e., within each one we can distinguish two principal stages: I – the calculations for the reference system and II – the EOM part. Stage I is identical to all EOM variants and may be formally divided into three steps: (i)solving Hartree–Fock equations via SCF calculation, (ii)solving coupled-cluster equation for the assumed reference state (iii)construction of the similarity transformed Hamiltonian, $\bar{H}_N$. Stage II relies on the diagonalization of the $\bar{H}_N$ matrix within the configurational subspace defined by the EOM variant. From a formal point of view, to recover the excited state of any atomic or molecular system (neutral closed-shell, radical, negative or positive ions) it is enough to have access to the EE-EOM-CC variant. By doing the reference calculation (Stage I) using general wave functions (RHF or UHF depending on the situation) we can apply the general purpose EE-EOM-CC scheme to get the EE states. However, in most cases we are forced to do a reference calculation for open-shell systems on the basis of UHF calculations with the well known limitation of that approach, e.g., spin contamination.

Having access to various realizations of the EOM-CC approach (EE, IP, EA, DIP, DEA, etc.) we propose a different computational strategy. In order to get access to excited states of species A we adopt one of the uncharged or charged analogues (A, A^+, A^-, A^{2+}, A^{2-}, …) as a reference which form a closed-shell system. Our experience tells us that even very sophisticated calculations are much easier to do for a RHF reference and to that aim we adjust our strategy. For instance, if we want to get the excited states of the A^- system we will use:

- the EE-EOM-CC method if A^- is a closed-shell species
- the EA-EOM-CC method if A turns out to be a closed-shell system
- the IP-EOM-CC method if A^{2-} turns out to be a closed-shell system.

Another example: in order to make successful calculations of potential energy curves for the AB molecule we assume one of the following systems: AB, AB^+, AB^-, AB^{2+}, AB^{2-} as reference which fulfill the condition that upon homolytic dissociation the closed-shell molecule creates closed-shell fragments. And according to this selection we will use either the EE-, EA-, IP-, DEA- or DIP-EOM-CC method, respectively.

For all five methods considered we have provided detailed working equations both at the CCSD and CCSDT level in the form ready to code. We are convinced that the computational method can be understood if we can see the working equations which can be directly coded into a computer program. For each of the considered EOM schemes we provided illustrative results which demonstrate the superiority of the EOM-CCSDT method. In several examples the CCSD errors are unacceptably large. In that case the obvious remedy is to include triple excitations.

We should mention that an efficient treatment of excited states is also possible within the EOM spin flip formalism of Krylov et al. [95–97] or the active space route [98, 99]. In EOM-sf the excited

states are sought in the basis of determinants conserving the total number of electrons but one has to change the number of α and β electrons. The reference function is computed for the high-spin state (e.g., S_z=1) then at the EOM level we arrive at the low-spin state (e.g., S_z=0) (see Ref. [55] for more details). A reverse spin-flip procedure has been applied recently in [100], where the spin non-conserving $\hat{R}(k)$ operator acting on the S_z=0 reference state produces the high-spin ($S_z \neq 0$) configurations. This made it possible to create a robust scheme to get triplets and quintets.

References

1 Coester, F. (1958). Bound states of a many-particle system. *Nucl. Phys.* 7: 421.
2 Coester, F. and Kümmel, H. (1960). Short-range correlation in nuclear wave functions. *Nucl. Phys.* 17: 477.
3 Čížek, J. (1966). On the correlation problem in atomic and molecular systems. Calculation of wave function components in Ursell-type expansion using quantum-field theoretical methods. *J. Chem. Phys.* 45: 4256.
4 Čížek, J. (1969). On the use of the cluster expansion and the technique of diagrams in calculations of correlation effects in atoms and molecules. *Adv. Chem. Phys.* 14: 15.
5 Paldus, J., Čížek, J., and Shavitt, I. (1974). Correlation problems in atomic and molecular systems. IV. Extended coupled-pair many-electron theory and its application to the BH_3 molecule. *Phys. Rev. A* 5: 50.
6 Purvis, G.D. and Bartlett, R.J. (1982). A full coupled-cluster singles and doubles model: the inclusion of disconnected triples. *J. Chem. Phys.* 76: 1910.
7 Bartlett, R.J. and Musiał, M. (2007). Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.* 79: 291.
8 Raghavachari, K., Trucks, G.W., Pople, J.A., and Head-Gordon, M. (1989). A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.* 157: 479.
9 Lyakh, D.I., Musiał, M., Lotrich, V., and Bartlett, R.J. (2012). Multireference nature of chemistry: the coupled-cluster view. *Chem. Rev.* 112: 182.
10 Rowe, D.J. (1968). The equations-of-motion method and the extended shell model. *Rev. Mod. Phys.* 40: 153.
11 Emrich, K. (1981). An extension of the coupled cluster formalism to excited states. *Nucl. Phys.* 351: 397.
12 Sekino, H. and Bartlett, R.J. (1984). A linear response, coupled-cluster theory for excitation energy. *Int. J. Quantum Chem.* 18: 255.
13 Geertsen, J., Rittby, M., and Bartlett, R.J. (1989). The equation-of-motion coupled-cluster method: excitation energies of Be and CO. *Chem. Phys. Lett.* 164: 57.
14 Comeau, D.C. and Bartlett, R.J. (1993). The equation-of-motion coupled-cluster method. Applications to open- and closed-shell reference states. *Chem. Phys. Lett.* 207: 414.
15 Stanton, J.F. and Bartlett, R.J. (1993). The equation of motion coupled-cluster method. A systematic biorthogonal approach to molecular excitation energies, transition probabilities, and excited state properties. *J. Chem. Phys.* 98: 7029.
16 Bartlett, R.J. and Stanton, J.F. (1994). *Reviews in Computational Chemistry*, vol. 5 (eds. K.B. Lipkowitz and D.B. Boyd), 65. New York: VCH Publishers.
17 Nooijen, M. and Bartlett, R.J. (1995). Equation of motion coupled cluster method for electron attachment. *J. Chem. Phys.* 102: 3629.
18 Watts, J.D. and Bartlett, R.J. (1994). The inclusion of connected triple excitations in the equation-of-motion coupled-cluster method. *J. Chem. Phys.* 101: 3073.

19 Watts, J.D. and Bartlett, R.J. (1995). Economical triple excitation equation-of-motion coupled-cluster methods for excitation energies. *Chem. Phys. Lett.* 233: 81.

20 Watts, J.D. and Bartlett, R.J. (1996). Iterative and non-iterative triple excitations in coupled-cluster methods for excited electronic states: the EOM-CCSDT-3 and EOM-CCSD ($\tilde{T}$) methods. *Chem. Phys. Lett.* 258: 581.

21 Musiał, M., Kucharski, S.A., and Bartlett, R.J. (2004). Approximate inclusion of the T3 and R3 operators in the equation-of-motion coupled cluster method. *Adv. Quantum Chem.* 47: 209.

22 Kowalski, K. and Piecuch, P. (2001). The active-space equation-of-motion coupled-cluster methods for excited electronic states: full EOMCCSDt. *J. Chem. Phys.* 115: 643.

23 Kucharski, S.A., Włoch, M., Musiał, M., and Bartlett, R.J. (2001). Coupled-cluster theory for excited states: the full equation-of-motion coupled-cluster single, double and triple excitation method. *J. Chem. Phys.* 115: 8263.

24 Hirata, S. (2004). Higher-order equation-of-motion coupled-cluster methods. *J. Chem. Phys.* 121: 51.

25 Kallay, M. and Surjan, P.R. (2000). Computing coupled-cluster wave functions with arbitrary excitations. *J. Chem. Phys.* 113: 1359.

26 Musiał, M., Kucharski, S.A., and Bartlett, R.J. (2003). Equation-of-motion coupled cluster method for the full inclusion of the connected triple excitations for ionized states: IP-EOM-CCSDT. *J. Chem. Phys.* 118: 1128.

27 Kamiya, M. and Hirata, S. (2006). Higher-order equation-of-motion coupled-cluster methods for ionization processes. *J. Chem. Phys.* 125: 074111.

28 Gour, J.R. and Piecuch, P. (2006). Efficient formulation and computer implementation of the active-space electron-attached and ionized equation-of-motion coupled-cluster methods. *J. Chem. Phys.* 115: 234107.

29 Musiał, M. and Bartlett, R.J. (2003). Equation-of-motion coupled cluster method with full inclusion of connected triple excitations for electron-attached states: EA-EOM-CCSDT. *J. Chem. Phys.* 119: 1901.

30 Kamiya, M. and Hirata, S. (2007). Higher-order equation-of-motion coupled-cluster methods for electron attachment. *J. Chem. Phys.* 126: 134112.

31 Sattelmeyer, K.W., Schaefer, H.F. III, and Stanton, J.F. (2003). Use of 2h and 3h – p-like coupled-cluster Tamm-Dancoff approaches for the equilibrium properties of ozone. *Chem. Phys. Lett.* 378: 42.

32 Demel, O., Shamasundar, K.R., Kong, L., and Nooijen, M. (2008). Application of double ionization state-specific equation of motion coupled cluster method to organic diradicals. *J. Phys. Chem. A* 112: 11895.

33 Musiał, M., Perera, A., and Bartlett, R.J. (2011). Multireference coupled-cluster theory: the easy way. *J. Chem. Phys.* 134: 114108.

34 Kuś, T. and Krylov, A.I. (2011). Using the charge stabilization technique in the double ionization potential equation-of-motion calculations with dianion references. *J. Chem. Phys.* 135: 084109.

35 Shen, J. and Piecuch, P. (2013). Doubly electron-attached and doubly ionized equation-of-motion coupled-cluster methods with 4-particle-2-hole and 4-hole-2-particle excitations and their active space extensions. *J. Chem. Phys.* 138: 194102.

36 Musiał, M., Kucharski, S.A., and Bartlett, R.J. (2011). Multireference double electron attached coupled cluster method with full inclusion of the connected triple excitations: MR-DA-CCSDT. *J. Chem. Theory Comput.* 7: 3088.

37 Monkhorst, H.J. (1977). Calculation of properties with the coupled-cluster method. *Int. J. Quantum Chem.* 11: 421.

38 Koch, H., Jensen, H.J.A., Jörgensen, P., and Helgaker, T. (1990). Excitation energies from the coupled cluster singles and doubles linear response function (CCSDLR). Applications to Be, CH^+, CO, and H_2O. *J. Chem. Phys.* 93: 3345.

39 Koch, H. and Jörgensen, P. (1990). Coupled cluster response functions. *J. Chem. Phys.* 93: 3333.

40 Koch, H., Christiansen, O., Jörgensen, P., and Olsen, J. (1995). Excitation energies of BH, CH_2 and Ne in full configuration interaction and the hierarchy CCS, CC2, CCSD and CC3 of coupled cluster models. *Chem. Phys. Lett.* 244: 75.

41 Christiansen, O., Koch, H., and Jörgensen, P. (1995). Response functions in the CC3 iterative triple excitation model. *J. Chem. Phys.* 103: 7429.

42 Nakatsuji, H. (1979). Cluster expansion of the wave function. Electron correlation in ground and excited states by SAC (symmetry-adapted-cluster) and SAC CI theories. *Chem. Phys. Lett.* 67: 329.

43 Nakatsuji, H. (1979). Cluster expansion of the wave function. Calculation of electron correlations in ground and excited states by SAC and SAC CI theories. *Chem. Phys. Lett.* 67: 334.

44 Nakatsuji, H., Ohta, K., and Hirao, K. (1981). Cluster expansion of the wave function. Electron correlation in ground state, valence and Rydberg excited states, ionized states, and electron-attached states of formaldehyde by SAC and SAC-CI theories. *J. Chem. Phys.* 75: 2952.

45 Noga, J. and Bartlett, R.J. (1988, 1987). The full CCSDT model for molecular electronic structure. *J. Chem. Phys.* 86: 7041. Errata: 89: 3401.

46 Kucharski, S.A. and Bartlett, R.J. (1992). The coupled-cluster single, double, triple, and quadruple excitation method. *J. Chem. Phys.* 97: 4282.

47 Kucharski, S.A. and Bartlett, R.J. (1999). Connected quadruples for the frequencies of O_3. *J. Chem. Phys.* 110: 8233.

48 Musiał, M., Kucharski, S.A., and Bartlett, R.J. (2002). Formulation and implementation of the full coupled-cluster method through pentuple excitations. *J. Chem. Phys.* 116: 4382.

49 Bartlett, R.J. (1989). Coupled-cluster approach to molecular structure and spectra: a step toward predictive quantum chemistry. *J. Phys. Chem.* 93: 1697.

50 Bartlett, R.J. (1995). *Modern Electronic Structure Theory, Part 2* (ed. D.R. Yarkony), 1047. New York: World Scientific Publishing.

51 Paldus, J. and Li, X. (1999). A Critical assessment of coupled cluster method in quantum chemistry. *Adv. Chem. Phys.* 110: 1.

52 Piecuch, P., Kucharski, S.A., Kowalski, K., and Musiał, M. (2002). Efficient computer implementation of the renormalized coupled cluster methods. The R-CCSD[T], R-CCSD(T), CR-CCSD[T] and CR-CCSD(T) approaches. *Comput. Phys. Commun.* 149: 71.

53 Piecuch, P., Kowalski, K., Pimienta, I.S.O. et al. (2004). Method of moments of coupled-cluster equations: a new formalism for designing accurate electronic structure methods for ground and excited states. *Theor. Chem. Acc.* 112: 349.

54 Shavitt, I. and Bartlett, R.J. (2009). *Many-Body Methods in Chemistry and Physics: Many-Body Perturbation Theory and Coupled Cluster Methods*. Cambridge, England: Cambridge Press.

55 Krylov, A.I. (2008). Equation-of-motion coupled-cluster methods for open-shell and electronically excited species: the Hitchhiker's guide to Fock space. *Annu. Rev. Phys. Chem.* 59: 433.

56 Musiał, M. (2010). The excited, ionized and electron attached states within the EOM-CC approach with full inclusion of connected triple excitations. *Mol. Phys.* 108: 2921.

57 Korona, T. and Werner, H.J. (2003). Local treatment of electron excitations in the EOM-CCSD method. *J. Chem. Phys.* 118: 3006.

58 Crawford, T.D. and King, R.A. (2002). Locally correlated equation-of-motion coupled cluster theory for the excited states of large molecules. *Chem. Phys. Lett.* 366: 611.

59 Nooijen, M. and Snijders, J.G. (1992). Coupled cluster approach to the single-particle Green's function. *Int. J. Quantum Chem.* 26: 55.

60 Nooijen, M. and Snijders, J.G. (1993). Coupled cluster Green's function method: working equations and applications. *Int. J. Quantum Chem.* 48: 15.

61 Stanton, J.F. and Gauss, J. (1994). Analytic energy derivatives for ionized states described by the equation-of-motion coupled cluster method. *J. Chem. Phys.* 101: 8938.

62 Stanton, J.F. and Gauss, J. (1999). A simple scheme for the direct calculation of ionization potentials with coupled-cluster theory that exploits established excitation energy methods. *J. Chem. Phys.* 111: 8785.

63 Szalay, P.G., Watson, T., Perera, A. et al. (2012). Benchmark studies on the building blocks of DNA. 1. Superiority of coupled cluster methods in describing the excited states of nucleobases in the Franck-Condon region. *J. Phys. Chem. A* 116: 6702.

64 Davidson, E.R. (1975). The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices. *J. Comput. Phys.* 17: 87.

65 Hirao, K. and Nakatsuji, H. (1982). A generalization of the Davidson's method to large nonsymmetric eigenvalue problems. *J. Comput. Phys.* 45: 246.

66 Musiał, M. and Bartlett, R.J. (2008). Intermediate Hamiltonian Fock-space multireference coupled-cluster method with full triples for calculation of excitation energies. *J. Chem. Phys.* 129: 044101.

67 Kendall, R.A., Dunning, T.H. Jr., and Harrison, R.J. (1992). Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* 96: 6796.

68 Li, X. and Paldus, J. (2006). General-model-space state-universal coupled-cluster methods for excited states: diagonal noniterative triple corrections. *J. Chem. Phys.* 124: 034112.

69 Musiał, M., Kucharski, S.A., Zerzucha, P. et al. (2009). Excited and ionized states of the ozone molecule with full triples coupled cluster methods. *J. Chem. Phys.* 131: 194101.

70 Sadlej, A.J. (1988). Medium-size polarized basis sets for high-level correlated calculations of molecular electric properties. *Collect. Czechoslov. Chem. Commun.* 53: 1995.

71 Banichevich, A. and Peyerimhoff, S.D. (1993). Theoretical study of the ground and excited states of ozone in its symmetric nuclear arrangement. *Chem. Phys.* 174: 93.

72 Anderson, S.M., Morton, J., and Mauersberger, K. (1990). Near-infrared absorption spectra of $^{16}O_3$ and $^{18}O_3$: adiabatic energy of the $_1A_2$ state? *J. Chem. Phys.* 93: 3826.

73 Thunemann, K.-H., Peyerimhoff, S.D., and Buenker, R.J. (1978). Configuration interactions calculations for the ground and excited states of ozone and its positive ion: energy locations and transitions probabilities. *J. Mol. Spectrosc.* 70: 432.

74 Kimura, K., Katsumata, S., Achiba, Y. et al. (1981). *Handbook of Hel Photoelectron Spectra of Fundamental Organic Molecules*. New York: Halsted.

75 Musiał, M., Bewicz, A., Skupin, P., and Kucharski, S.A. (2018). Potential energy curves for the LiK^+ and NaK^+ molecular ions with the coupled cluster method. *Adv. Quantum Chem.* 76: 333.

76 Radziemski, L.J., Engelman, R. Jr., and Brault, J.W. (1995). Fourier-transform-spectroscopy measurements in the spectra of neutral lithium, ^{6}I and ^{7}I (Li I). *Phys. Rev. A* 52: 4462.

77 Sansonetti, J.E. (2008). Wavelengths, transition probabilities, and energy levels for the spectra of sodium (Na I-Na XI). *J. Phys. Chem. Ref. Data* 37: 1659.

78 Musiał, M. and Bartlett, R.J. (2011). Multi-reference Fock space coupled-cluster method in the intermediate Hamiltonian formulation for potential energy surfaces. *J. Chem. Phys.* 135: 044121.

79 Siegbahn, H., Asplund, L., and Kelfve, P. (1975). The Auger electron spectrum of water vapour. *Chem. Phys. Lett.* 35: 330.

80 Thompson, M., Hewritt, P.A., and Wooliscroft, D.S. (1976). Analytical aspects of Auger electron spectrometry of gases. *Anal. Chem.* 48: 1336.

81 Rye, R.R., Madey, T.E., Houston, J.E., and Holloway, P.H. (1978). Chemical-state effects in Auger electron spectroscopy. *J. Chem. Phys.* 69: 1504.

82 Moddeman, W.E., Carlson, T.A., Krause, M.O. et al. (1971). Determination of the *K-LL* Auger spectra of N_2, O_2 CO, NO, H_2O, and CO_2. *J. Chem. Phys.* 55: 2317.

83 Dawber, G., McConkey, A.G., Acaldi, L. et al. (1994). Threshold photoelectrons coincidence spectroscopy of doubly-charged ions of nitrogen, carbon monoxide, nitric oxide and oxygen. *J. Phys. B Atomic Mol. Phys.* 27: 2191.

84 Benoit, C. and Horsley, J.A. (1975). Electronic states of $C_2H_4^{++}$ from double charge transfer spectroscopy and from SCF-CI calculations. *Mol. Phys.* 30: 557.

85 Musiał, M. (2012). Multi-reference Fock space coupled-cluster method in standard and intermediate Hamiltonian formulation for the (2,0) sector. *J. Chem. Phys.* 136: 134111.

86 Moore, C.E. (1949). Atomic energy levels. *Natl. Bur. Stand. (U.S.)* 467: 60.

87 Koch, H., Kobayashi, R., Sanchez de Meras, A., and Jörgensen, P. (1994). Calculation of size-intensive transition moments from the coupled cluster singles and doubles linear response function. *J. Chem. Phys.* 100: 4393.

88 Mukhopadhyay, D., Kukhopadhyay, S., Chaudhuri, R., and Mukherjee, D. (1991). Aspects of separability in the coupled cluster based direct methods for energy differences. *Theor. Chem. Acc.* 80: 441.

89 Stanton, J.F. (1994). Separability properties of reduced and effective density matrices in the equation-of-motion coupled cluster method. *J. Chem. Phys.* 101: 8928.

90 Musiał, M. and Bartlett, R.J. (2011). Charge-transfer separability and size-extensivity in the equation-of-motion coupled cluster method: EOM-CC. *J. Chem. Phys.* 134: 034106.

91 Nooijen, M. and Bartlett, R.J. (1997). A new method for excited states: similarity transformed equation-of-motion coupled-cluster theory. *J. Chem. Phys.* 106: 6441.

92 Nooijen, M. and Bartlett, R.J. (1997). Similarity transformed equation-of-motion coupled-cluster study of ionized, electron attached, and excited states of free base porphin. *J. Chem. Phys.* 106: 6449.

93 Meissner, L. (1998). Fock-space coupled-cluster method in the intermediate Hamiltonian formulation: model with singles and doubles. *J. Chem. Phys.* 108: 9227.

94 Musiał, M. and Bartlett, R.J. (2008). Multireference Fock-space coupled-cluster and equation-of-motion coupled-cluster theories: the detailed interconnections. *J. Chem. Phys.* 129: 134105.

95 Krylov, A.I. (2001). Size-consistent wave functions for bond-breaking: the equation-of-motion spin-flip model. *Chem. Phys. Lett.* 338: 375.

96 Levchenko, S.V. and Krylov, A.I. (2004). Equation-of-motion spin-flip coupled-cluster model with single and double substitutions: theory and application to cyclobutadiene. *J. Chem. Phys.* 120: 175.

97 Levchenko, S.V. and Krylov, A.I. (2005). Spin-conserving and spin-flipping equation-of-motion coupled-cluster method with triple excitations. *J. Chem. Phys.* 123: 84107.

98 Kowalski, K. and Piecuch, P. (2000). the active-space equation-of-motion coupled-cluster methods for excited electronic states: the EOMCCSDt approach. *J. Chem. Phys.* 113: 8490.

99 Krylov, A.I., Sherrill, C.D., and Head-Gordon, M. (2000). Excited states theory for optimized orbitals and valence optimized orbitals coupled-cluster doubles models. *J. Chem. Phys.* 113: 6509.

100 Musiał, M., Lupa, Ł., and Kucharski, S.A. (2016). Equation-of-motion coupled cluster method for the description of the high spin excited states. *J. Chem. Phys.* 144: 154105.

5

The Algebraic-Diagrammatic Construction Scheme for the Polarization Propagator

Andreas Dreuw

Interdisziplinäres Zentrum für Wissenschaftliches Rechnen, Ruprecht-Karls Universität Heidelberg, Im Neuenheimer Feld 205, 69120 Heidelberg, Germany

Abstract

The algebraic-diagrammatic construction scheme (ADC) for the polarization propagator is introduced and the key steps of the original derivation via Green's functions as well as a more recent one via the intermediate state representation (ISR) are presented. The ISR provides an elegant and efficient route to excited state properties and excited-state wave function analysis tools. Properties and limitations of the ADC scheme for excitation energies are discussed, before specialized variants of ADC like unrestricted ADC (UADC), spin-flip (SF) ADC, spin-opposite-scaled (SOS) ADC and core-valence-separated (CVS) ADC are described in a bit more detail. The chapter concludes with the description of photochemistry by ADC and a brief summary and outlook.

Over the last ten years, the algebraic diagrammatic construction scheme for the polarization propagator of second order (ADC(2)) has become more and more popular as the excited-state method for the investigation of photochemistry of medium-sized, mostly organic molecules [1, 2]. It is also often employed as the benchmark method for computationally more efficient but also more approximate excited-state methods like time-dependent density functional theory or semi-empirical methods. It can be used as an alternative to coupled-cluster approaches. The success of ADC(2) is certainly related to the ongoing efforts to provide user-friendly and efficient computer codes. Meanwhile, the ADC(2) method is implemented in a whole bunch of quantum chemical program packages, for example, in Turbomole [3], Psi4 [4], Orca [5], and Q-Chem [6] to name a few. In addition, ADC approaches are so-called black-box methods and offer user-friendly application because expert knowledge is not required to run the calculation. In other words, no *a priori* knowledge of the molecular system is required and only a basis set and the number of excited states to be calculated have to be chosen to run an ADC calculation. As *ab initio* methods, ADC approaches possess a stable and predictable and widely established accuracy for a large class of molecules.

However, ADC(2) is only one specialized approach of a whole family of ADC methods and as will be shown below, algebraic diagrammatic construction (ADC) refers to a general formalism to construct computational schemes starting from Green's functions, i.e., propagators. It was developed in the early 1980s for the calculation of ionization potentials, electron affinities and excitation energies [1, 7, 8]. In analogy to coupled-cluster approaches it is thus important to indicate the variant by either specifying the propagator that has been constructed via the ADC

Quantum Chemistry and Dynamics of Excited States: Methods and Applications,
First Edition. Edited by Leticia González and Roland Lindh.

formalism or by specifying the property to be computed. ADC(n) for the polarization propagator yields excitation energies (EE), while ADC(n) for the hole or electron propagator yields ionization potentials (IP) or electron affinities (EA). Hence, it is suggested that the established coupled-cluster nomenclature is followed, i.e., to use the abbreviation EE-ADC, IP-ADC and EA-ADC to specify the variant of ADC. Since, in the following, the focus lies on EE-ADC schemes exclusively, here ADC(n) will always refer to EE-ADC(n).

This chapter is organized as follows. Following the historical pathway, the original derivation of ADC(n) schemes via Green's function, i.e., propagators, will be first briefly reviewed (Section 5.1) before the approach of the intermediate state representation (ISR) is introduced (Section 5.2), and exploited to compute excited-state and transition properties (Section 5.3.1). In Section 5.4, connections to other theoretical methods are made and properties and limitations of EE-ADC schemes are discussed. Afterwards specialized variants of EE-ADC are introduced (Section 5.5), before finally, in Section 5.6, the simulation of organic photochemistry with EE-ADC approaches is described. The chapter concludes with a brief summary and outlook of future developments.

5.1 Original Derivation via Green's Functions

The Green's function formalism represents a successful technique to solve inhomogeneous differential equations [9]. Instead of solving the inhomogeneous differential equation, if it is possible at all, a corresponding differential equation is solved, in which the inhomogeneity $I(r)$ is replaced by the Dirac delta distribution $\delta(r - r')$. The solution of this modified differential equation is called *Green's function*. Knowing the Green's function, the solution of the full inhomogenous differential equation can then be obtained via

$$f(r) = \int dr' G(r - r') I(r'). \tag{5.1}$$

In analogy, a Green's function can be defined for the time-dependent single-particle Schrödinger equation

$$\left\{ \imath \frac{\partial}{\partial t} - \hat{H}(r) \right\} G(rt; r't') = \delta(r - r')\delta(t - t') \tag{5.2}$$

This is particularly useful if the Hamiltonian $\hat{H}$ can be separated into one part $\hat{H}_0$ for which the Schrödinger equation can be solved and a small perturbation $\hat{V}(r)$. Then, the solution of the Schrödinger equation with $\hat{H}$ is given by

$$|\Psi(rt)\rangle = |\Psi_0(rt)\rangle + \int dr' \int dt' G(rt; r't') \hat{V}(r') |\Psi_0(r't')\rangle, \tag{5.3}$$

where $|\Psi_0(rt)\rangle$ is the solution of the Schrödinger equation with $\hat{H}_0$.

In many-body systems the definition of a Green's function as in the single-particle case is generally no longer possible. However, similar building blocks can be defined which yield solutions to certain problems within the many-body system. These building blocks are also called Green's functions or *propagators*. Most important are the one- and two-particle Green's function, which are defined as [10]

$$G_{p,q}(t_p, t_q) = -\imath \langle \Psi_0 | \hat{T} \left\{ \hat{a}_p(t_p) \hat{a}_q^\dagger(t_q) \right\} | \Psi_0 \rangle \tag{5.4}$$

$$G_{p,q;r,s}(t_p, t_q; t_r, t_s) = (-\imath)^2 \langle \Psi_0 | \hat{T} \left\{ \hat{a}_p(t_p) \hat{a}_q(t_q) \hat{a}_r^\dagger(t_r) \hat{a}_s^\dagger(t_s) \right\} | \Psi_0 \rangle \tag{5.5}$$

where $|\Psi_0\rangle$ refers to the exact N-particle ground state. Many-body Green's functions are ground-state expectation values of a time-ordered series of creation ($\hat{a}_p^\dagger(t)$) and annihilation ($\hat{a}_p(t)$) operators, which create or annihilate a particle at time t in a single-particle orbital $\phi_p(r)$. The single-particle Green's function thus gives the probability that an electron that has been in the single-particle state q at time t_q appears in the single-particle state p at later time t_p, and the two-particle Green's function analogously for two electrons.

Since Green's functions are defined as ground-state expectation values, some information about the ground state of the system is lost compared to the ground state wave function. However, the information content of the one-particle Green's function Eqn. (5.4) still suffices to calculate ground-state expectation values of single-particle operators, the ground-state energy, as well as the ionization and absorption spectra of the system. The two-particle Green's function Eqn. (5.5) can be employed to calculate ground-state expectation values of two-particle operators and the excitation spectrum. For the latter, even the *polarization propagator* is sufficient, which is obtained from the two-particle Green's function via

$$\Pi_{pq,rs}(t,t') = \lim_{t_q \to t} \lim_{t_s \to t'} \imath \{ G_{p,s;q,r}(t,t_s;t_q,t') - G_{p,q}(t,t_q) G_{s,r}(t_s,t') \}. \tag{5.6}$$

The first term on the right hand side of Eqn. (5.6) corresponds to the two-particle Green's function related to the probability that at time t two electrons in the single-particle states p and s occur at time t' in the states q and r. The product $G_{p,q}(t,t_q)G_{s,r}(t_s,t')$ is subtracted, which describes the probability that the single-particle transitions p to q and s to r occur independently from each other. Hence, the difference corresponds to the influence of one single-particle transition onto the other. Physically this influence is *polarization*, which becomes clear when both fluctuations are imagined to occur on spatially separated atoms for example. If the atoms do not interact, density fluctuations occur independently, and only when the atoms polarize each other, density fluctuations are vice versa induced, and the *polarization propagator* is non-zero.

A more intuitive expression for the polarization propagator is obtained after insertion of a compete set of eigenfunctions and Fourier transform into energy space. This so-called spectral Lehmann representation takes on the form

$$\begin{aligned} \Pi_{pq,rs}(\omega) &= \sum_{(n\neq 0)} \frac{\langle \Psi_0 | \hat{a}_q^\dagger \hat{a}_p | \Psi_n \rangle \langle \Psi_n | \hat{a}_r^\dagger \hat{a}_s | \Psi_0 \rangle}{\omega + E_0 - E_n} \\ &\quad + \sum_{(n\neq 0)} \frac{\langle \Psi_0 | \hat{a}_r^\dagger \hat{a}_s | \Psi_n \rangle \langle \Psi_n | \hat{a}_q^\dagger \hat{a}_p | \Psi_0 \rangle}{-\omega + E_0 - E_n} \qquad (5.7) \\ &= \Pi^+_{pq,rs}(\omega) + \Pi^-_{pq,rs}(\omega). \qquad (5.8) \end{aligned}$$

Π^+ has poles at excitation energies and Π^- carries information about "de-excitations" of the system. For the derivation of the second- or third-order polarization propagator approaches, SOPPA and TOPPA [11, 12], respectively, the full polarization propagator is taken into account. In contrast, in the algebraic diagrammatic construction schemes only Π^+ is considered, which in its diagonal form, i.e., in the eigenstate basis of the Hamiltonian, reads as

$$\Pi^+(\omega) = \mathbf{x}^\dagger (\omega - \mathbf{\Omega})^{-1} \mathbf{x}. \tag{5.9}$$

For the derivation of approximate expressions for the polarization propagator the existence of a non-diagonal form is assumed

$$\Pi^+(\omega) = \mathbf{f}^\dagger (\omega - \mathbf{M})^{-1} \mathbf{f}, \tag{5.10}$$

and *diagrammatic* perturbation theory is used to *construct algebraic* expressions for the matrix elements of **f** and **M**

$$\mathbf{M} = \mathbf{K} + \mathbf{C}^{(1)} + \mathbf{C}^{(2)} + \cdots \ \mathbf{C}^{(n)} \tag{5.11}$$

$$\mathbf{f} = \mathbf{f}^{(0)} + \mathbf{f}^{(1)} + \mathbf{f}^{(2)} \ \cdots \ \mathbf{f}^{(n)} \tag{5.12}$$

Without going further into the mathematical details, which can be found in Ref. [1], the order n of perturbation theory used for the matrix elements of **M**, the ADC matrix, and **f**, the matrix of modified transition moments, determines the order of the algebraic diagrammatic construction scheme ADC(n). More specifically, n refers to the order of perturbation theory up to which the primary particle–hole (p–h) states are exact. **K** corresponds in Eqn. (5.11) to zeroth order, i.e., the expectation value of the Fock operator of the zeroth-order intermediate states, the Slater determinants. Hence, **K** is diagonal with orbital energy difference between occupied and virtual orbitals. Diagonalization of the matrix **M**, i.e., solution of the corresponding Hermitian eigenvalue equation, yields directly excitation energies ω_n and ADC vectors $\mathbf{y}_n$ according to

$$\mathbf{MY} = \mathbf{\Omega Y} \quad \text{with} \quad \mathbf{Y}^\dagger\mathbf{Y} = \mathbf{1}, \tag{5.13}$$

with **Y** and Ω being the matrices of ADC eigenvectors and excitation energies, respectively. The transition moments are then obtained from the converged ADC vectors via

$$\mathbf{x}_n = \mathbf{y}_n^\dagger \mathbf{f}, \tag{5.14}$$

i.e., by contraction with the previously determined modified transition moments (Eqn. (5.12)).

The unusual name of Algebraic Diagrammatic Construction originates from this first derivation via diagrammatic perturbation theory to construct algebraic expressions and it refers to a general mathematical procedure and is generally applicable to any propagator. As we have seen, ADC for the polarization propagator yields a Hermitian eigenvalue equation with excitation energies and transition vectors as eigenvalues and eigenvectors. When the one-particle propagator is used (Eqn. 5.4), Hermitian ADC schemes are obtained which yield ionization potentials or electron affinities. Hence, it is important to refer to the propagator used to classify the ADC scheme. For example, ADC(3) for the polarization propagator yields excitation energies of primary particle–hole states consistently up to third order [13, 14], while ADC(3) for the hole-propagator yields ionization potentials [7, 15]. In analogy to coupled-cluster approaches, the abbreviations EE-ADC, IP-ADC or EA-ADC for excitation energy, ionization potential or electron affinity ADC are thus suggested to be used in a more general context.

5.2 The Intermediate State Representation

An alternative approach for the derivation of ADC schemes is provided by the formalism of the intermediate state (IS) representation [16, 17]. In the original derivation via the Green's function formalism, the existence of a non-diagonal form of the polarization propagator (Eqn. 5.10) has been postulated. Almost 20 years later, the basis of the non-diagonal form leading to the ADC expressions for the polarization propagator has been identified. This is the so-called *Intermediate State (IS)* basis.

Conceptually, the construction of the IS basis is simple. By acting with physical excitation operators on the correlated electronic ground state according to

$$|\Psi_I^{CE}\rangle = \hat{C}_I|\Psi_0\rangle \quad \text{with} \quad \hat{C}_I = \{\hat{a}_a^\dagger\hat{a}_i; \hat{a}_b^\dagger\hat{a}_j\hat{a}_a^\dagger\hat{a}_i \ \cdots\}, \tag{5.15}$$

a set of correlated excited-state basis functions is first constructed. $\hat{a}_a^\dagger$ and $\hat{a}_i$ correspond to the typical creation and annihilation operators of second quantization. The Ψ_I^{CE} are generally not orthogonal and are thus step-wise orthogonalized with respect to the electron ground state and other excitation classes. This results in the orthogonal intermediate state basis Ψ_I^{IS} for excited states. When this basis is used to represent the shifted Hamiltonian $(\hat{H} - E_0)$, with E_0 being the correlated ground state energy, a matrix is obtained which reads

$$\{\mathbf{M}\}_{IJ} = \langle \Psi_I^{IS} | \hat{H} - E_0 | \Psi_J^{IS} \rangle. \tag{5.16}$$

Clearly, diagonalization of $\mathbf{M}$, i.e., solving the eigenvalue equation $\mathbf{MY} = \mathbf{\Omega Y}$ yields excitation energies. So far, this derivation is exact as long as the exact ground state wave function and ground state energy are employed in Eqn. (5.15) and (5.16), respectively. For the derivation of ADC schemes, however, the Møller–Plesset ground state energy as well as the corresponding wave function are employed. In general, the perturbation theoretical order of the ADC(n) scheme is then determined by the order of the MPn quantities used. Thereby, identical expressions for the matrix elements of $\mathbf{M}$ are obtained as via the traditional route of diagrammatic perturbation theory. In other words, the IS basis is that basis, in which the ADC matrix and so the converged ADC vectors $\mathbf{y}_n$ are eventually represented. Therefore, knowledge of the IS basis vectors $\{\Psi_I^{IS}\}$ is required to construct an excited state wave function from an ADC eigenvector via

$$\Psi_n = \sum_I y_{nI} \Psi_I^{IS}, \tag{5.17}$$

where I runs over all intermediate states, the ground state excluded. ADC amplitudes are important to realize – not to correspond to coefficients of Slater-determinants as in configuration interaction – but to expansion coefficients of a correlated excited state basis. Hence, an interpretation of ADC vectors in the orbital picture is not appropriate. Equation (5.17) provides furthermore access to linear and non-linear excited-state properties at ADC level by computing expectation values or evaluating higher-order response functions (*vide infra*).

It is worthwhile to note that using MP0 values in the ISR formalism, i.e., the HF ground state and the expectation value of the Fock-operator $\sum_i \epsilon_i$, yields an ADC(0) matrix with the dimension $n_{occ} \times n_{virt}$ with just orbital energy differences $\epsilon_a - \epsilon_i$ on the diagonal. Using MP1 values, ADC(1) is obtained which is identical to the uncorrelated CIS method. Only when MP2 or MP3 serve as a starting point for the ISR derivation, ADC(2) or ADC(3), respectively, are obtained, which are

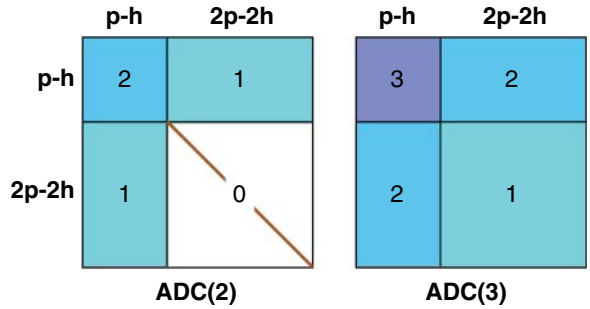


Figure 5.1 Structure of the ADC(2) and ADC(3) matrices. The number in the corresponding block refers to the order of perturbation theory to which the matrix elements of the block are exact. For example, the (p–h/p–h) block of the ADC(2) matrix is exact up to second order, the (p–h/2p-2h) and (2p-2h/p–h) coupling blocks up to first order and the (2p-2h/2p-2h) block only up to zeroth order. In other words the latter block has only orbital energy differences $(\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j)$ on the diagonal.

correlated excited state methods. Both matrices exhibit the dimension of a CISD matrix containing singly-excited, also called particle–hole (p–h) states and doubly excited, two-particle two-hole (2p-2h) states (Figure 5.1). Until today, the ADC(4) scheme for the polarization propagator has not been realized, and its ADC(4) matrix would extend to the (3p-3h) space, i.e., have the dimension of a CISDT matrix.

Within the derivation of the ADC schemes the MPn ground state energy is *formally* subtracted, canceling the corresponding terms in the Hamiltonian $\hat{H} = \hat{F} + \hat{U}$, which is perturbation theoretically expanded exactly. Hence these ground state terms do not occur in the ADC matrix, and thus a ground state MP calculation is generally not required prior to an ADC calculation for excitation energies only. The only necessary ingredients are one-particle energies and molecular orbitals stemming from a (canonical) Hartree–Fock calculation. This is important to realize, as it will become relevant when external one-particle potentials are to be included in an ADC calculation (Section 5.6.2).

Finally it should be briefly mentioned that besides the derivation via Green's functions and the ISR approach, ADC equations can also be obtained via further routes. For ADC(2), yet another alternative route exists via the configuration interaction singles with perturbative doubles correction (CIS(D_∞) scheme [18]. It has been recognized that the ADC(2) matrix is obtained, when the Lagrangian of the CIS(D_∞) is symmetrized according to

$$\mathbf{M}_{ADC(2)} = \frac{1}{2}\{\mathbf{A}_{CIS(D_\infty)} + \mathbf{A}^{\dagger}_{CIS(D_\infty)}\}. \tag{5.18}$$

In the context of CIS(D_∞) or CC2, one speaks of a Lagrangian instead of a Hamiltonian, as these matrices are not Hermitian. Since CC2 is related to CIS(D_∞) by a similarity transformation of the Hamiltonian, CC2 is also closely related to ADC(2) explaining their similar accuracies and range of applicability. More details on this derivation route can be found in the literature [18].

Early on, it was realized that unitary coupled cluster (UCC) provides yet a different derivation route for ADC schemes in general [19, 20], if appropriate assumptions are made during the derivation. Without going into detail here, one of the key assumptions lies in the termination of the Baker–Campbell–Hausdorff expansion, which does not terminate naturally for the UCC ansatz. However, it has also been shown recently that UCC polarization propagator theory can lead to Hermitian UCC3 equations for excited states which are structurally identical to ADC(3) [21]. The difference lies in the description of the electronic ground state. While at ADC(3) level it is simply the MP3 ground state, UCC(3) requires the solution of the ground-state UCC3 equation. Hence, UCC3 is computationally more expensive than ADC(3), for cases where MP3 yields a poor ground state description, UCC3 performs slightly better than ADC(3).

5.3 Calculation of Excited State Properties and Analysis

5.3.1 Excited State Properties

The calculation of transition properties as well as excited-state properties, such as transition dipole moments between excited states or the dipole moment of an excited state, is straightforwardly possible by virtue of the ISR (Section 5.2) [16, 17]. Diagonalization of the ADC matrix yields the corresponding ADC vectors $\mathbf{y}_n$, which are the expansion coefficients of excited state wave functions represented in the IS basis according to Eqn. (5.17). For the computation of an expectation value of excited state n

$$\mu_n = \langle \Psi_n | \hat{\mu} | \Psi_n \rangle \tag{5.19}$$

the complete IS basis can be inserted

$$\mu_n = \sum_I \langle \Psi_n | \Psi_I^{IS} \rangle \langle \Psi_I^{IS} | \hat{\mu} | \Psi_I^{IS} \rangle \langle \Psi_I^{IS} | \Psi_n \rangle \tag{5.20}$$

yielding

$$\mu_n = \mathbf{y}_n^\dagger \mathbf{D}^{IS} \mathbf{y}_n, \tag{5.21}$$

where $\mathbf{D}^{IS}$ is the representation of the dipole operator in the IS basis. Hence, for the computation of an excited-state dipole moment as expectation value, the dipole operator needs to be represented in the IS basis and the resulting matrix contracted with the corresponding ADC vector, or equivalently, with the ADC excited-state density. Just as simple, the transition dipole moment μ_{nm} between two excited states can be calculated as

$$\mu_{nm} = \mathbf{y}_n^\dagger \mathbf{D}^{IS} \mathbf{y}_m. \tag{5.22}$$

Since the expressions for the IS basis are analytically known, the formalism is not only limited to the dipole operator, but can easily be extended to other first-order properties. It has already been successfully applied to compute different excited-state and transition properties, for example, spin–orbit coupling elements within the AMFI approximation [22].

Not only can first-order properties easily be computed as expectation values using the ISR formalism, but also higher-order response properties like non-linear optical properties are accessible [23–25]. For example, resonant two-photon absorption cross sections for the simultaneous absorption of two photons given as a sum-over-states expression as

$$S_{\alpha\beta} = \sum_n^\infty \left(\frac{\langle \Psi_0 | \hat{\mu}^\alpha | \Psi_n \rangle \langle \Psi_n | \hat{\mu}^\beta | \Psi_f \rangle}{\omega_n - \frac{\omega}{2}} + \alpha \leftrightarrow \beta \right), \tag{5.23}$$

with $\alpha, \beta \in \{x, y, z\}$. Using the ISR, Eqn. (5.23) is transformed into a matrix expression

$$S_{\alpha\beta} = \mathbf{F}_\alpha^\dagger (\mathbf{M} - \frac{\omega}{2})^{-1} \mathbf{B}_\beta \mathbf{Y_f} + \frac{\mu_{00}^\alpha \mu_{0f}^\beta}{E_0 \frac{\omega}{2}} + \alpha \leftrightarrow \beta, \tag{5.24}$$

where $\mathbf{F}_\alpha$ is the matrix of modified transition moments along the direction α, $\mathbf{B}_\beta$ is the IS representation of the dipole operator in direction β, and $\mathbf{Y}_f$ is the ADC eigenvector of the final state. It should be noted that the intermediate state basis does not contain the ground state and hence the ground state contribution is not included in the matrix representation and needs to be added separately. In principle, the ADC matrix has to be inverted and thus fully diagonalized to determine the two-photon cross section, which is computationally quite demanding. To circumvent this costly procedure, a vector $\mathbf{x}_\alpha^\dagger$ is defined

$$\mathbf{x}_\alpha^\dagger = \mathbf{F}_\alpha^\dagger \left(\mathbf{M} - \frac{\omega}{2} \right)^{-1}, \tag{5.25}$$

which is thus the solution of the following set of linear equations

$$\mathbf{x}_\alpha^\dagger \left(\mathbf{M} - \frac{\omega}{2} \right) = \mathbf{F}_\alpha^\dagger. \tag{5.26}$$

Once this system of linear equations is solved for each component $\alpha \in \{x, y, z\}$, pseudo-densities are built

$$\rho^{\alpha f} = \mathbf{x}_\alpha^\dagger \mathbf{D} \mathbf{Y}_f, \tag{5.27}$$

which are eventually used to compute the two-photon cross sections according to

$$S_{\alpha\beta} = \sum_{pq} \rho^{\alpha f} \mu_{pq}^\beta + \frac{\mu_{00}^\alpha \mu_{0f}^\beta}{E_0 - \frac{\omega}{2}} + \alpha \leftrightarrow \beta. \tag{5.28}$$

Within the ISR formalism outlined so far, excited state properties are computed as expectation values of the corresponding wave function, which according to the postulates of quantum mechanics is the formally rigorous approach. However, excited-state properties can also be computed as derivative of the excited state energy. For example, the dipole moment of an excited state is also defined as the derivative of the total energy with respect to the electric field components at zero field strength.

$$\mu_n^\alpha = \left(\frac{\partial E_n}{\partial F_\alpha}\right)_{F=0} \quad \text{mit} \quad \alpha \in \{x, y, z\}. \tag{5.29}$$

Accordingly, the electric dipole moment and all excited-state properties that can be formulated as derivative of the energy can be computed following this procedure. In modern quantum chemistry codes, these derivatives are calculated using a Lagrangian formalism [3–6]. This gives rise to the distinction between relaxed and unrelaxed properties, i.e., whether orbital relaxation is taken into account or not, respectively [26]. The calculation of relaxed excited-state properties is computationally more expensive, since a coupled-perturbed Hartree–Fock equation needs to be solved for each excited state to determine the orbital relaxation. Of course, the calculated excited state properties generally differ depending on whether the ISR formalism, or relaxed or unrelaxed derivatives are employed, however, in the limit of an exact theory, all approaches converge to the same value.

5.3.2 Excited-State Wave Function and Density Analyses

As mentioned previously, the expansion coefficients of the ADC vector are inappropriate to be interpreted in the uncorrelated picture of HF orbitals as they refer to correlated IS basis functions according to Eqn. (5.17). The key for any analysis of the electronic structure of excited states at ADC level is, however, also only possible by virtue of Eqn. (5.17) making a wave function accessible at all. Also for density-based analyses of the electronic structure of excited states in ADC, knowledge of the ISR is indispensable for the construction of the excited-state one-electron densities and one-electron transition densities (1-TDM). The latter represent an excellent means to visualize and thus understand the electronic excitation process as the 1-TDM reveals the *vertical* nature of the electronic transition [27–29]. It is defined as

$$T_{pq}^{0n} = \langle \Psi_0 | \hat{a}_p^\dagger \hat{a}_q | \Psi_n \rangle. \tag{5.30}$$

The single-particle functions diagonalizing the 1-TDM are the natural transition orbitals (NTO) [30], which can be, and often are, utilized to visualize the nature of the corresponding electronic transition. Associated with the 1-TDM, the so-called electron-hole amplitude of the polarization propagator can be defined as [28]

$$\chi(r_h, r_e) = \sum_{p,q} T_{pq}^{0n} \phi_p^*(r_h) \phi_q(r_e), \tag{5.31}$$

recognizing the application of $\hat{a}_p^\dagger$ creates a hole in orbital ϕ_p whereas the operator $\hat{a}_q$ acts to the right to annihilate an electron in orbital ϕ_q. The two-particle wave function $\chi(r_h, r_e)$ can be interpreted as an exciton wave function nicely connecting the exciton picture with molecular excited electronic states [31]. Based on this exciton wave function and among other things, hole and particle densities can be assigned, as

$$\rho_h(r_h) = \int |\chi(r_h, r_e)|^2 d^3 r_e, \tag{5.32}$$

and

$$\rho_e(r_e) = \int |\chi(r_h, r_e)|^2 d^3 r_h, \tag{5.33}$$

respectively. Plotting these densities nicely reveals the nature of the electron transitions.

A complementary approach to understand the electronic structure of an excited state and not the electronic transition is provided by the one-electron difference density matrix (1-DDM), which is the difference between electron densities of the ground and excited states. [28, 29] It is thus defined as

$$\Delta_{pq} = P^n_{pq} - P^0_{pq} = \langle \Psi_n | \hat{a}^\dagger_p \hat{a}_q | \Psi_n \rangle - \langle \Psi_0 | \hat{a}^\dagger_p \hat{a}_q | \Psi_0 \rangle. \tag{5.34}$$

The corresponding one-electron difference density is given by

$$\rho_\Delta(r) = \sum_{p,q} \Delta_{pq} \phi^*_p(r) \phi_q(r), \tag{5.35}$$

which integrates to zero and can already be plotted to visualize the excited state electronic structure. However, the difference density possesses positive and negative parts and sometimes a complicated nodal structure. In such cases, the so-called detachment/attachment density plots offer a more convenient way for analyzing the electronic structure [32]. For their construction, Δ is first diagonalized and then decomposed into its negative and positive parts, the diagonal detachment **D** and attachment **A** density matrices, respectively. After re-transformation into the molecular orbital basis, the attachment density is obtained as

$$\rho_A(r) = \sum_{p,q} A_{pq} \phi^*_p(r) \phi_q(r). \tag{5.36}$$

and can be easily plotted. The detachment density is obtained analogously.

Summarizing, hole/electron plots characterize the vertical nature of the electronic transition while detachment/attachment density plots show the electronic structure of the final state. Their comparison can nicely visualize orbital relaxation processes occurring in excited states, which are particularly pronounced, for instance, in core-excited states [33]. Since the amplitudes of the ADC vector do not refer to Slater-determinants, but to correlated excited-state basis functions, a direct interpretation in the molecular orbital is not adequate and only zeroth-order. Therefore, one should always resort to density-based analyses when interpreting ADC results, depending on the problem and personal taste: transition or difference-density based approaches.

5.4 Properties and Limitations of ADC

In general, all EE-ADC schemes are represented by Hermitian eigenvalue equations (Eqn. (5.13)). Hence, there exists only one set of eigenvectors, and not like in standard equation-of-motion (EOM) or linear response (LR) CC schemes left and right ones, which are generally both required to rigorously compute transition moments and properties. Solving ADC(2) and ADC(3) equations thus boils down to a straightforward diagonalization of the corresponding matrix using some iterative diagonalization scheme, like the Davidson or Lanczos procedure [34, 35]. In the case of ADC(2), a DIIS formalism can be employed and the contributions of the diagonal (2p-2h/2p-2h)-block can be folded into the (p–h) block, such that the dimension of the matrix is reduced to that of a CIS matrix [36].

In contrast to configuration interaction schemes, ADC total energies given as $E_I(ADC(n)) = E_0(MPn) + \omega_I(ADC(n))$ are size-extensive, i.e., they scale linearly with the number of independent particles, because the excitation energies are size-intensive and the MP2 ground state

Table 5.1 Comparison of the general properties of single-reference configuration interaction (CI), coupled-cluster (CC) and ADC methods.

	CI	CC	ADC
hermiticity	yes	no	yes
size-consistency	no	(yes)[a)]	yes
compactness	n	m+1/m+2[b)]	m+1
ground state	CI	CC	MP

a) Energies are always size-extensive, properties and transition moments only in LR-CC
b) for even/odd order of perturbation theory (see text).

energy is size-extensive [1]. The excitation energies are independent of the number of computed non-interacting molecules, as they are strictly localized on one molecule only. Also the transition moments and other excited-state properties are size-intensive at ADC level of theory. Comparing to EOM-CC and LR-CC approaches, the energies are also size-extensive. The transition moments and properties are size-intensive in case of LR-CC but not in case of EOM-CC.

Compared to CI and CC schemes, ADC is more compact. In general, compactness of an excited state method is defined as the extension of the configuration space required to describe primary excited (p–h) states correct to a certain order n in perturbation theory. For example, doubly excited states are required in CI to describe (p–h) states correct to second order, triply excited states for a correct description to third order. Hence, the compactness of standard single-reference CI is linear, n. For even numbered orders of perturbation theory $n = 2m$, ADC and CC are equivalently compact, as both need the (2p-2h) space to describe (p–h) states correct in second order perturbation theory, and their compactness is $m + 1$. For odd numbered orders $n = m + 1$, ADC is more compact than CC. While for example CC theories require already (3p-3h) configurations to describe (p–h) states correct up to third order, ADC(3) stays within the (2p-2h) configuration space. Hence, the compactness for odd-numbered order amounts to $m + 2$ for CC theories and only $m + 1$ for ADC schemes. This manifests itself also in the computational effort. While all second-order EOM- or LR-CC schemes and ADC(2) scale formally as $O(N^5)$, with N being the number of single-particle basis functions, ADC(3) scales only as $O(N^6)$ in contrast to CC3 ($O(N^7)$) or EOM/LR-CCSDT ($O(N^8)$) [2, 37].

Until the derivation route via the ISR was found, the electronic ground state was neither accessible nor rigorously known. Only by virtue of the ISR, has MPn been identified as the most obvious reference state. This perturbative MP reference state is also the largest drawback of ADC methods, which limits standard ADC methods to molecules which are already fairly well described by a single Slater determinant, i.e., via Hartree–Fock theory. This limits the applicability range of ADC essentially to well-behaved organic molecules, i.e., organic photochemistry. Closed-shell metal complexes can also be described [38]. In addition, as will be described below in more detail, MP2 is usually only applicable in the vicinity of the equilibrium geometry, and as soon as large distortions or bond breaking occurs, the ground state usually acquires multi-reference character and the MP description breaks down. The ADC also becomes no longer reliable and large errors in the excitation energy are likely to occur. CC approaches offer a better ground state treatment, via the corresponding underlying ground-state coupled-cluster equations. Hence, their range of applicability is

slightly larger and the methods remain stable longer when approaching multi-reference situations. However, prior to EOM/LR-CC excitation energy calculations the ground-state equation must be solved, which is not the case in ADC, adding some extra computational cost to CC methods. As we will see below, tricks exist to extend the range of single-reference ADC and CC methods alike when multi-reference situations are encountered via the so-called spin-flip approach (Section 5.5.3).

Finally, let us have a look at the accuracy of the standard closed-shell EE-ADC schemes. Since they are only applicable to single-reference systems, for which MP2 yields an accurate ground-state description, the appropriate test set is the one devised by Thiel [39, 40], which provided theoretical best estimates for 141 vertical excited singlet and 71 triplet states of 28 small to medium-sized organic molecules. The standard ADC(2) scheme has revealed good accuracies of 0.22±0.38 eV for singlet and 0.12±0.16 eV for triplet excited states of this test set, which is very similar to the accuracy of CC2, which however has a smaller standard deviation [14]. Going one order higher in perturbation theory to ADC(3), the accuracy gets higher. With an error of only 0.12±0.28 eV for singlet states and −0.18±0.16 eV for triplet states, ADC(3) is highly accurate and can be safely used as a benchmark method for lower level methods, provided medium-sized closed-shell organic molecules are the target [14]. A recent benchmark on a set of 18 very small molecules with up to three non-hydrogen atoms revealed larger mean absolute errors of ADC(3) of 0.28 and 0.17 eV for valence and Rydberg states, respectively [41]. Third-order CC methods perform substantially better for this subset, which may be related to their improved ground state description and larger importance of dynamical correlation in smaller molecules. In the context of ADC methods, the so-called Dyson expansion method allows to further include higher-order terms in the self-energy Σ_∞ and the ground-state one-particle density matrix to improve the description of the electronic ground state and thus the overall performance of ADC(3). However, this has not yet been demonstrated and remains to be numerically tested.

5.5 Variants of EE-ADC

The traditional derivation route of ADC schemes requires the existence of an appropriate propagator. This prevented, for a long time, the derivation of specialized variants of EE-ADC schemes, for example like unrestricted ADC, since no open-shell polarization propagator could be defined. In contrast, the derivation route via the construction of an ISR basis provides this flexibility. In fact, the choice of the MPn scheme as well as the definition of the set of annihilation and creation operators in Eqn. (5.15) allow for the derivation of different variants of EE-ADC, which will be demonstrated in the following. It is important to note that in general the formal limitations of ADC apply to all variants as well. However, due to the similar overall structure of the variants, analysis tools and property calculations are available and readily applicable to all EE-ADC variants within Q-Chem.

5.5.1 Extended ADC(2)

In the ADC(2) scheme [42], primary (p–h) states are correct up to second order perturbation theory, while (2p-2h) states are given only at zeroth order leading to a tremendous energetic overestimation of (2p-2h) states, i.e., too large excitation energies. In an attempt to improve their description within ADC(2), the extended ADC(2) scheme (ADC(2)-x) has been devised, in which the matrix elements of the (2p-2h) block are *ad hoc* extended to first order in perturbation theory without any formal justification [42]. This extension leads to an increased computational effort from $O(N^5)$ for ADC(2) to now $O(N^6)$ for ADC(2)-x. Indeed, the excitation energies of primarily (2p-2h)

states are drastically reduced, but this *ad hoc* extension leads to an imbalanced description of the valence-excited states and to a substantial underestimation of the excitation energies with an overall lower accuracy than that of the parent ADC(2) method. The extended ADC(2)-x scheme has an overall low accuracy of only -0.70 ± 0.37 eV for singlet-excited states and -0.55 ± 0.20 for triplet states. It is thus fair to say, that ADC(2)-x is not a useful excited-state method for the investigation of the low-lying valence-excited states of molecular systems.

However, ADC(2)-x has its value for the calculation of core-excited states, i.e., X-ray absorption spectra, as will be seen below (Section 5.5.5), and as a diagnostic tool for the importance of doubly-excited states by comparison of results obtained with ADC(2) and ADC(2)-x [43]. Typically, the excitation energies of valence excited (p–h) states are lowered by a few tenths of an eV when going from ADC(2) to ADC(2)-x, while excitation energies of states with large double-excitation character are lowered by several eV. For example, the singly-excited $1B_u$ state of s-*trans*-butadiene is obtained at 7.09 eV and 6.69 eV at ADC(2) and ADC(2)-x level when the 6-31G basis set is used, while the doubly-excited $2A_g$ state occurs at 8.34 and 5.19 eV, respectively [43]. Such a strong lowering of the excitation energy as in the latter case is indicative of doubly excited states, and excludes ADC(2) and related second-order approaches as useful excited-state methods for the system under investigation. Instead one needs to resort to some higher level of theory like for instance ADC(3), EOM-CCSDT, CC3 or multi-reference methods.

5.5.2 Unrestricted EE-ADC Schemes

Although an open-shell version of the polarization propagator is missing, unrestricted EE-ADC schemes can be formally derived via the intermediate state representation following Eqn. (5.15). Instead of using a closed-shell reference MP ground state, the corresponding unrestricted MP scheme can serve as a starting point for the construction of the ISR and the shifted Hamiltonian-matrix $\mathbf{M}$ [44]. In this case, the annihilation and creation operators in Eqn. (5.15) must refer to spin-orbitals

$$\hat{C}_I^U = \{\hat{a}_{a\sigma}^\dagger \hat{a}_{i\sigma}; \hat{a}_{b\tau}^\dagger \hat{a}_{j\tau} \hat{a}_{a\sigma}^\dagger \hat{a}_{i\sigma} \ \cdots\}, \tag{5.37}$$

instead of spatial orbitals as in the closed-shell reference case. Following the same ISR derivation route to arrive at UADC(2), for example, $\hat{C}_I^U$ acts on the unrestricted UMP2 ground state and an unrestricted IS basis $\{\Psi_I^{UIS}\}$ is obtained, in which the shifted Hamiltonian is represented

$$\{\mathbf{M}\}_{IJ}^{UADC(2)} = \langle \Psi_I^{UIS} | \hat{H} - E_0^{UMP2} | \Psi_J^{UIS} \rangle. \tag{5.38}$$

Again, it is in principle not required to perform an UMP2 calculation prior to the UADC(2) calculation, since only the unrestricted set of molecular orbitals and corresponding orbital energies of an underlying UHF calculation are needed to build the matrix elements of the UADC matrix during the iterative diagonalization. However, UHF can suffer from spin-contamination, which has a general influence also on the accuracy of UADC. Therefore it is useful to perform an UMP2 calculation, anyways, since its accuracy and applicability is a prerequisite for the UADC schemes to give reliable results. In other words, if UMP2 works fine for the molecular open-shell system of interest, UADC(2) and UADC(3) will work well too.

A pathway to circumvent spin-contamination is to start the ISR derivation from a restricted open-shell reference instead of an unrestricted reference. However, this has not yet been realized, essentially for two main reasons. On one hand, it requires changes in the block structure of the ADC matrix, as unoccupied, singly- and doubly-occupied orbitals need to be distinguished and the corresponding matrix elements evaluated. One the other hand, and more importantly, it seems

more advisable to follow an IP-ADC route to excited states of open-shell systems, in strict analogy to IP-EOM-CC schemes [45, 46]. Here, the calculation starts from a closed-shell (2n)-electronic reference and ionization potentials and open-shell (2n-1)-electronic states are calculated. The excitation spectrum of the open-shell species can then be obtained as differences between the lowest calculated IP_0 and the higher IP_n, ω_n=IP_n-IP_0, and the transition dipole moments by virtue of the corresponding IP-ISR representation of the dipole operator in analogy to Eqn.(5.22). This procedure naturally yields spin-pure open-shell states, and as in the case of IP-EOM-CC schemes probably also to higher accuracy.

5.5.3 Spin-Flip EE-ADC Schemes

In standard ADC schemes, restricted or unrestricted, the singlet MP ground-state is used as reference and the excitation operators $\hat{C}_J$ are restricted to $\Delta m_s = 0$ excitations to conserve the spin of the system (Eqns. (5.15) and (5.37)). Therefore, standard ADC schemes are expected to yield accurate and reliable results for molecular systems with single-reference electronic ground states, which are well described with Møller–Plesset perturbation theory. This limitation essentially excludes the description of systems with multi-reference character including rare cases like open-shell singlets, like diradical or diradicaloid systems, but also important chemical situations like single-bond dissociation and ground-state conical intersections.

A clever trick to extend the range of applicability of single-reference methods also to a certain set of multi-reference situations is the spin-flip idea [47, 48]. In spin-flip methods, the triplet ground state is used as reference as it often is a single-reference state even when the singlet ground state has multi-reference character. This is for example typically the case in single-bond breaking situations. Therefore, in spin-flip ADC schemes [49], the triplet ground state is chosen as reference and the excitation operators need to be restricted to perform a spin flip (Δm_s=−1) to generate all basis functions, including the singlet ground state as well as all triplet states with m_s=0 according to

$$\hat{C}_I^{SF} = \{\hat{a}^\dagger_{a\beta}\hat{a}_{i\alpha}; \hat{a}^\dagger_{b\tau}\hat{a}_{j\tau}\hat{a}^\dagger_{a\beta}\hat{a}_{i\alpha} \ \cdots\}, \tag{5.39}$$

In the (2p-2h) excitations, only one spin is flipped, while the second excitation conserves spin and follows the Δms=0 rule. Using the resulting IS basis for the representation of the shifted Hamiltonian readily yields SF-ADC schemes

$$\mathbf{M}^{SF}\mathbf{X} = \mathbf{\Omega}\boldsymbol{X} \quad \text{with} \quad \mathbf{X}^\dagger\mathbf{X} = 1. \tag{5.40}$$

Straightforward diagonalization of $\mathbf{M}^{SF}$ yields excitation energies of the target states relative to the triplet ground state and the corresponding ADC transition vectors in the IS basis. The target states comprise “all” relevant singlet and triplet excited states, including the potentially multi-reference singlet ground state. As a consequence, SF-ADC schemes allow for the investigation of diradicaloid molecules with open-shell singlet ground states, single-bond breaking and ground-state conical interactions [50]. Moreover, SF-ADC schemes yield highly accurate results for singlet-triplet gaps [51], because singlets and triplets are treated on an absolutely identical footing, i.e., in the reference space of the SF-ADC matrix. In analogy to UADC approaches, SF-ADC approaches require only a preceding UHF calculation. Therefore, the calculated target states are generally not spin-pure and can be spin contaminated leading to a tedious discrimination of singlet and triplet states. In contrast to UADC, SF-ADC approaches can also easily make use of a restricted open-shell HF reference calculation reducing the spin-contamination. However, in all tested cases this had however only little influence on the accuracy of the SF-ADC results [49].

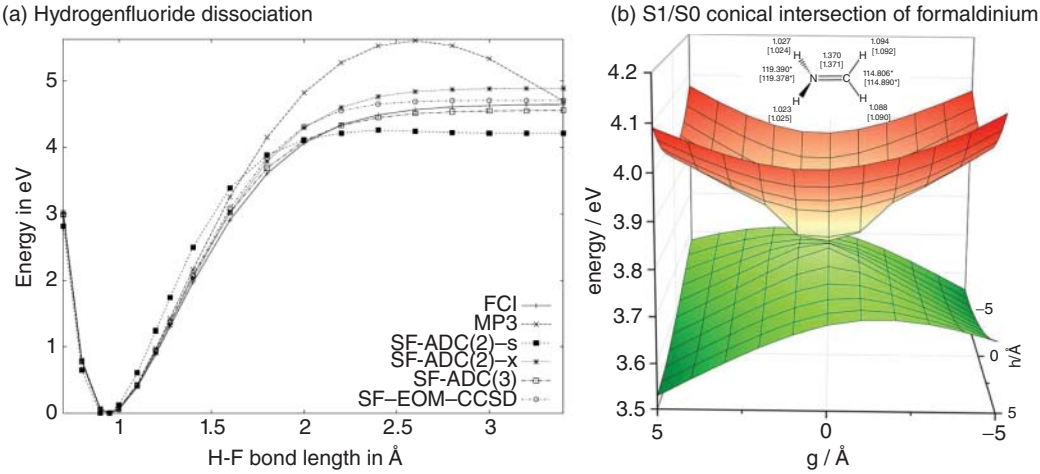


Figure 5.2 (a) Ground state potential energy curves of hydrogen fluoride at various levels of SF-ADC. For comparison also the MP3, full-CI and SF-EOM-CCSD curves are shown. (b) Optimized S_1/S_0 conical intersection of formaldinium at the level of SF-ADC(3). The optimized minimum energy structure is shown in the inset and compared to CASPT2 values [in brackets].

To demonstrate the applicability of SF-ADC approaches to multi-reference situations, the potential energy curves of the ground state of hydrogen fluoride (HF) obtained at various levels of SF-ADC and the optimized conical intersection of formaldinium at the level of SF-ADC(3) are shown in Figure 5.2. While the parent ground state method, i.e., MP3 experiences a variational break down along the dissociation of HF all SF approaches are capable of obtaining the qualitatively correct shape of the dissociation curve. Of course, SF-ADC(3) agrees better with the full configuration interaction (full-CI) curve than ADC(2)-x and ADC(2). The optimized minimum energy conical intersection at SF-ADC(3) level agrees remarkably well with the one obtained at CASPT2 level. Moreover, also the potential energy surface along the coupling and tuning mode, i.e., the cone of the intersection are in excellent agreement (Figure 5.2(b)). At this point it is important to realize the SF-ADC(3) calculations to be straightforward black-box calculations without the need for tedious choice of active space or other parameters. The plot in Figure 5.2(b) has been obtained by simple scanning along the relevant modes only choosing the method and basis set.

However, a final word of warning must be noted. SF-ADC methods are not genuine multi-reference methods as the system under investigation is required to possess a stable single-reference triplet state for the methods to provide reliable results. This is typically the case in situations when the singlet ground state is dominated by two leading references, in other words, when two electrons are decoupled as in single-bond breaking or double-bond rotation. For such cases the term “few-reference” situation has been coined, and therefore, all SF methods should thus be termed rather “few-reference” methods and not “multi-reference”.

5.5.4 Spin-Opposite-Scaled ADC Schemes

To reduce the computational effort and to increase the accuracy of MP2, spin-component scaling (SCS) and spin-opposite scaling (SOS) variants of MP2 have been developed, in which the same-spin components and opposite-spin components of the correlation energy are individually scaled with semi-empirical factors [52, 53]. Among the different variants, SOS-MP2 is particularly interesting,

since here only the spin-opposite correlation energy is calculated and scaled appropriately

$$E_{SOS-MP2} = c_{os} E^{os}_{MP2} \tag{5.41}$$

with $c_{os} = 1.3$ reducing the cost of an MP2 calculation. The SOS-MP2 ground state energy and corresponding wave function can of course serve as a basis for the derivation of SOS-ADC(2) schemes following the ISR formalism as outlined above. However, a rigorous derivation leads to an ISR-SOS-ADC(2) scheme of the same size as the original ADC(2), because the only change occurs in a scaling of the t-amplitudes in the (ph/ph)-block. This substitution leads to a systematic shift of the excitation energies by about 0.02–0.3 eV to lower but not improved excitation energies [53]. At the same time, no significant saving in computational effort can be achieved as no entries of the original ADC(2)-matrix vanish.

To arrive at a computationally more efficient SOS-ADC(2) scheme, further approximations have to be made and additional semi-empirical parameters need to be introduced. In analogy to SOS-CC2, an efficient SOS-ADC(2) scheme can be devised by: (1) replacing all t-amplitudes by their spin-opposite scaled analogs as in SOS-MP2 (and ISR-SOS-ADC(2)), (2) deleting all entries from same spin contributions in the double excitation manifold and (3) scaling of all remaining entries in the (ph/2p2h) and (2p2h/ph) coupling blocks by the additional factor of $c^{os}_{coupling}$, which can be freely chosen as for example

$$c^{os}_{coupling} = \frac{1 + c_{os}}{2} = 1.15, \tag{5.42}$$

or be fitted to minimize the deviation from benchmark values. Extensive fitting revealed c_{os}=1.3 and $c^{os}_{coupling}$=1.17 to possess the smallest errors in singlet and triplet excitation energies compared to theoretical best estimates for organic molecules with a mean absolute error of only 0.14±0.11 eV [53]. The deletion of same-spin (2p2h) configurations leads in addition to a significant reduction of the size of the ADC matrix and a reduction of the computational effort.

The concept of semi-empirical scaling has been extended to ADC(2)-x as well to devise a second-order method providing a balanced description of singly- and doubly-excited state. Therefore a third scaling factor c_x is introduced in the off-diagonal elements of (2p-2h) block thus modulating the coupling strength between the (2p-2h) configurations. Fitting against theoretical best estimates and DFT/MRCI values revealed optimal values of c_{os}=1.3, $c^{os}_{coupling}$=1 and C_x=0.9 for SOS-ADC(2)-x. Statistical error analysis revealed an error of this semi-empirical method of 0.17±0.16 eV for singly-excited states and 0.21eV±0.18 eV for doubly-excited states [53]. This SOS-ADC(2)-x is a pragmatic semi-empirical approximation to ADC(2)-x reducing its computational cost substantially. However, prior to ist use for the investigation of organic photochemistry a thorough testing and benchmarking of the quality of the results against higher-order methods is strongly recommended.

5.5.5 Core-Valence Separated (CVS) EE-ADC

Electronic excitations of a core-electron into the valence region, so-called core-excited states, represent a special class of excited states. Their calculation with standard excited-state methods is usually difficult, because the electronic spectrum of a molecule is usually converged from the low-energy end towards higher energies. Also, the ADC eigenvalue problem is solved numerically via iterative diagonalization schemes yielding the energetically lowest excited states. Since core-excited states are in the high energy X-ray region of the spectrum, this solution strategy becomes cumbersome and quite expensive due to the necessity to compute all energetically lower-lying excited states

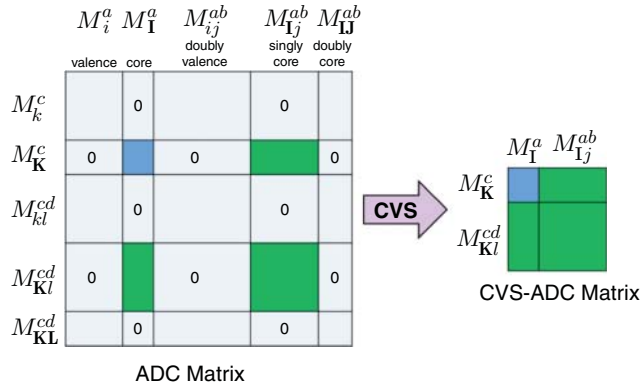


Figure 5.3 By neglecting all couplings of the singly core-excited states to the valence-excited and doubly core-excited states, the CVS-ADC matrix is obtained which contains singly core-excited states only.

beforehand. This would limit the calculation of core-excited states to very small molecular systems. With a simple trick, however, it is possible to avoid this problem, i.e., the ADC scheme can be combined with the core-valence separation (CVS) approximation [54].

The CVS approximation exploits the fact that core-excited states lie in a very different energy regime than valence-excited states and couple thus only very weakly, i.e., practically not, to them. In the CVS approximation, these coupling are consequently set to zero.

$$\langle Ip|qr\rangle = \langle pI|qr\rangle = \langle pq|Ir\rangle = \langle pq|rI\rangle = 0, \tag{5.43}$$

$$\langle IJ|pq\rangle = \langle pq|IJ\rangle = 0, \tag{5.44}$$

$$\langle IJ|Kp\rangle = \langle IJ|pK\rangle = \langle Ip|JK\rangle = \langle pI|JK\rangle = 0. \tag{5.45}$$

Here, I, J, K refer to core-orbitals while p, q, r refer to valence orbitals. This procedure leads to a strict decoupling of the singly core-excited states from the valence-excited and the doubly core-excited configurations. As a consequence, the singly core-excited block, the CVS-ADC matrix (Figure 5.3), can be diagonalized individually, yielding the core-excited states directly. Thereby also significant savings in computational effort are achieved. Instead of *a posteriori* decoupling of the core-excitations from the rest, the CVS-ADC matrix is similarly obtained via the ISR route, if in Eqn. (5.15) exactly one annihilation operator of each excitation class is chosen to correspond to a core-orbital.

In the context of EE-ADC schemes, the CVS approximation has been applied to the restricted and unrestricted versions of ADC(2), ADC(2)-x as well as ADC(3) [55–57]. For a selected set of small molecules, CVS-ADC(3) has been shown to possess a mean error of 0.61%±0.32% at the CBS limit. In contrast, CVS-ADC(2)-x exhibits a smaller error of only −0.23%±0.1% [57]. Using the 6-311++G** basis set in combination with CVS-ADC(2)-x yields even better results with a very small error of only about 0.1%. It has been shown that this is, of course, due to fortuitous error compensation of basis set truncation, neglect of relativistic effects, electron correlation, and orbital relaxation. In the meantime, many K-edge X-ray absorption spectra of organic molecules have been computed using CVS-ADC(2)-x/6-311++G**, and the accuracy has always been that high proving the described error compensation to be stable for typical organic molecules with second-row elements. This error compensation is then broken at the third order level, because the ratio between relaxation and polarization effects is changed and the excitation energy increases [57]. Core-excited

state properties and densities can be computed in strict analogy to valence-excited states following the ISR formalism as described above in Section 5.3.1 [58].

5.6 Describing Molecular Photochemistry with ADC Methods

In principle, the theoretical description of molecular photochemistry, i.e., photoinduced processes in molecules, requires the treatment of quantum or semi-classical dynamics of nuclear motion, since their outcome is practically always kinetically controlled and not, as in ground state chemistry, thermodynamically. In other words, the involved energy barriers do not necessarily determine the course of the reaction, in particular when they are fast, but much more an interplay of the initial forces driving the reaction into a particular reaction, the amount of excess energy and their dissipation as well as competing radiative processes determine the result of this time-scale competition. However, dynamical simulations are often simply not feasible either because the molecule is too large or the number of required nuclear degrees of freedom is not tractable. In such cases, "static" investigations of the potential energy surfaces (PES) of the electronic states involved are often the only manageable pathway to gain insights into the ongoing mechanisms. Recently, ADC methods have more and more been used to perform such investigations and are used as benchmark methods to evaluate the accuracy of computationally more efficient but less reliable approaches like semi-empirical methods or time-dependent density functional theory. In addition, PES, nuclear excited-state gradients, non-adiabatic coupling elements, spin-orbital couplings, transition moments, etc. are anyways necessary ingredients for dynamical simulations. In this context, ADC(2) has already been used as underlying electronic structure method to devise these required informations [59].

5.6.1 Potential Energy Surfaces

With ADC being originally a propagator method, yielding "just" excitation energies and neither total energies nor excited-state wave functions, all the above mentioned properties, first and foremost potential energy surfaces, could not be rigorously calculated. Usually a ground state potential surface at some suitable level of theory (often even at the Hartree–Fock level) was chosen and the excitation energies added to obtain excited-state surfaces. Only with the advent of the ISR formalism, MPn has been clarified to be the correct reference ground state method for ADC(n). Hence, total energies of excited states are obtained as

$$E_n^{ADC(n)} = E_0^{MPn} + \omega_n^{ADC(n)}, \tag{5.46}$$

i.e., the sum of MPn ground state energy and ADC(n) excitation energy [16]. Although this allows for a more rigorous construction of excited-state PES, they inherit the problems of the underlying ground state MPn method. This includes, for example, the variational collapse of MPn along single-bond breaking (as seen in Figure 5.2), and one can generally not expect MP approaches to remain accurate farther away from the ground-state equilibrium geometry, where the ground state usually acquires multi- or few-reference character as described in section 5.5.3. Therefore, whenever ADC methods are used as underlying electronic structure method for nuclear dynamics simulations their applicability must constantly be carefully checked. However, if ADC is applicable, all other required ingredients, like excited-state gradients, non-adiabatic couplings or spin-orbit couplings are available [59].

5.6.2 Environment Models within ADC

Another important aspect for the theoretical investigation of photochemistry and, in particular, for the simulation of spectroscopy of medium-sized and large molecules, is the inclusion of the molecular environment in the calculation. For most experimental studies are performed in solution or molecular matrices, or the chromophore of interest is embedded in a protein environment. Very scarcely, an experimental spectroscopic investigation on larger organic compounds is performed in the gas phase.

In the context of EE-ADC, the inclusion of molecular environments, which are often represented as one-particle potentials, is pretty straightforward and here only the general strategy will be presented rather than individual environment models discussed. Due to the structure of the ADC matrix (Eqn. (5.11)), two different pathways for treating external one-particle potentials $v_{ext}(r)$ within ADC are generally possible. On one hand, it can be included in the Fock-operator, $\hat{F} = \hat{F}_0 + v_{ext}(r)$ whit $\hat{F}_0$ being the unperturbed Fock operator, of the preceding Hartree–Fock calculation to obtain perturbed one-particle energies and orbitals. Subsequently, a standard ADC calculation is performed utilizing those. On the other hand, the one-particle potential can be added to the ADC matrix after an unperturbed HF calculation has been performed

$$\mathbf{M} = \mathbf{M}_0 + \mathbf{V}_{ext}, \tag{5.47}$$

with $\mathbf{M}_0$ being the unperturbed ADC matrix and $\mathbf{V}_{ext}$ is the external potential represented in the IS basis. At approximate ADC levels, both approaches offer advantages, however, to include a one-particle potential at HF level provides fully relaxed orbitals with respect to the potential and only minimal changes to the existing ADC codes need to be made. The latter is thus the route usually followed in the development of environment models within ADC. State-specific corrections for the solvent influence are then usually performed *a posteriori*, i.e., after the ADC calculation has finished, as perturbative corrections. Following this procedure, polarizable continuum models (PCMs) [60–62], polarizable embedding (PE) [63], the effective fragment potential (EFP) method [64], and frozen-density embedding (FDE) [65, 66] have been combined with EE-ADC methods and their accuracy established.

5.7 Brief Summary and Perspective

The gist of the ADC family method is best described by a quote from the very first introductory paper by Jochen Schirmer: "Thus, the ADC may be viewed as representing an advantageous *mixtum compositum* of a diagonalization problem and perturbation theory." [1]. Since then ADC methods have matured to valuable computational tools for organic photochemistry, as they have been implemented into several quantum chemistry packages and they are enjoying increasing popularity in various fields of application. Their reliability, accuracy and range of applicability are nowadays well established. ADC schemes augment existing established excited state methods, possessing different strengths and weaknesses than others, thereby complementing the toolbox of the computational photochemist.

In the future, the limits of the ADC methods need to be addressed, which are practically all connected to the poor description of the electronic ground state by plain Møller–Plesset perturbation theory. One promising avenue to improve and to achieve a higher accuracy in vertical excitation energies, may be the approach via unitary CC theory as briefly sketched in section 5.2. However, this will mostly not solve the problem with multi-reference situations like bond-breaking or

ground-state conical intersections, the single-reference UCC ground state will probably just break later. Another avenue for future development in this direction would be to follow a multi-reference path, similar to multi-reference coupled cluster, however, this would mean sacrificing computational efficiency as well as the black-box character of ADC, two appealing features responsible for its current success.

Bibliography

1 Schirmer, J. (1982). Beyond the random-phase approximation: A new approximation scheme for the polarization propagator. *Phys. Rev. A* 26 (5): 2395–2416.
2 Dreuw, A. and Wormit, M. (2015). The algebraic diagrammatic construction scheme for the polarization propagator for the calculation of excited states. *WIREs Comput. Mol. Sci.* 5 (1): 82–95.
3 Furche, F., Ahlrichs, R., Hättig, C. et al. (2014). Turbomole. *WIREs Comput. Mol. Sci.* 4 (2): 91–100.
4 Turney, J.M., Simmonett, A.C., Parrish, R.M. et al. (2012). Psi4: an open-source ab initio electronic structure program. *WIREs Comput. Mol. Sci.* 2 (4): 556–565.
5 Neese, F. (2012). The ORCA program system. *WIREs Comput. Mol. Sci.* 2 (1): 73–78.
6 Krylov, A.I. and Gill, P.M. (2013). Q-Chem: an engine for innovation. *WIREs Comput. Mol. Sci.* 3 (3): 317–326.
7 Schirmer, J., Cederbaum, L.S., and Walter, O. (1983). New approach to the one-particle Green's function for finite Fermi systems. *Phys. Rev. A* 28 (3): 1237–1259.
8 Tarantelli, A. and Cederbaum, L.S. (1989). Particle-particle propagator in the algebraic diagrammatic construction scheme at third order. *Phys. Rev. A* 39 (4): 1656–1664.
9 Hazewinkel, M. (ed.) (1994). *Encyclopedia of Mathematics*. Netherlands: Springer.
10 Schirmer, J. (2018). *Many-Body Methods for Atoms, Molecules and Clusters*. Springer Berlin Heidelberg.
11 Packer, M.J., Dalskov, E.K., Enevoldsen, T. et al. (1996). A new implementation of the second-order polarization propagator approximation (SOPPA): The excitation spectra of benzene and naphthalene. *J. Chem. Phys.* 105 (14): 5886–5900.
12 Oddershede, J. (1978). Polarization Propagator Calculations. *Adv. Quantum Chem.*: 275–352.
13 Trofimov, A.B., Stelter, G., and Schirmer, J. (1999). A consistent third-order propagator method for electronic excitation. *J. Chem. Phys.* 111 (22): 9982–9999.
14 Harbach, P.H.P., Wormit, M., and Dreuw, A. (2014). The third-order algebraic diagrammatic construction method (ADC(3)) for the polarization propagator for closed-shell molecules: Efficient implementation and benchmarking. *J. Chem. Phys.* 141 (6): 064113.
15 Schirmer, J., Trofimov, A.B., and Stelter, G. (1998). A non-Dyson third-order approximation scheme for the electron propagator. *J. Chem. Phys.* 109 (12): 4734–4744.
16 Schirmer, J. and Trofimov, A.B. (2004). Intermediate state representation approach to physical properties of electronically excited molecules. *J. Chem. Phys.* 120 (24): 11 449–11 464.
17 Trofimov, A., Krivdina, I., Weller, J., and Schirmer, J. (2006). Algebraic-diagrammatic construction propagator approach to molecular response properties. *Chem. Phys.* 329 (1–3): 1–10.

18 Hättig, C. (2005). Structure optimizations for excited states with correlated second-order methods: CC2 and ADC(2). *Adv. Quantum Chem.*: 37–60.

19 Mertins, F. and Schirmer, J. (1996). Algebraic propagator approaches and intermediate-state representations. I. The biorthogonal and unitary coupled-cluster methods. *Phys. Rev. A* 53 (4): 2140–2152.

20 Schütz, M. (2015). Oscillator strengths, first-order properties, and nuclear gradients for local ADC(2). *J. Chem. Phys.* 142 (21): 214103.

21 Liu, J., Asthana, A., Cheng, L., and Mukherjee, D. (2018). Unitary coupled-cluster based self-consistent polarization propagator theory: a third-order formulation and pilot applications. *J. Chem. Phys.* 148 (24): 244110.

22 Krauter, C.M., Schimmelpfennig, B., Pernpointner, M., and Dreuw, A. (2017). Algebraic diagrammatic construction for the polarization propagator with spin-orbit coupling. *Chem. Phys.* 482: 286–293.

23 Knippenberg, S., Rehn, D.R., Wormit, M. et al. (2012). Calculations of nonlinear response properties using the intermediate state representation and the algebraic-diagrammatic construction polarization propagator approach: two-photon absorption spectra. *J. Chem. Phys.* 136 (6): 064107.

24 Knippenberg, S., Gieseking, R.L., Rehn, D.R. et al. (2016). Benchmarking Post-Hartree-Fock methods to describe the nonlinear optical properties of polymethines: an investigation of the accuracy of Algebraic Diagrammatic Construction (ADC) approaches. *J. Chem. Theory Comput.* 12 (11): 5465–5476.

25 Rehn, D.R., Dreuw, A., and Norman, P. (2017). Resonant inelastic X-ray scattering amplitudes and cross sections in the Algebraic Diagrammatic Construction/Intermediate State Representation (ADC/ISR) approach. *J. Chem. Theory Comput.* 13 (11): 5552–5559.

26 Salter, E.A., Sekino, H., and Bartlett, R.J. (1987). Property evaluation and orbital relaxation in coupled cluster methods. *J. Chem. Phys.* 87 (1): 502–509.

27 Tretiak, S. and Mukamel, S. (2002). Density matrix analysis and simulation of electronic excitations in conjugated and aggregated molecules. *Chem. Rev.* 102 (9): 3171–3212.

28 Plasser, F., Wormit, M., and Dreuw, A. (2014). New tools for the systematic analysis and visualization of electronic excitations. I. Formalism. *J. Chem. Phys.* 141 (2): 024106.

29 Plasser, F., Bäppler, S.A., Wormit, M., and Dreuw, A. (2014). New tools for the systematic analysis and visualization of electronic excitations. II. Applications. *J. Chem. Phys.* 141 (2): 024107.

30 Martin, R.L. (2003). Natural transition orbitals. *J. Chem. Phys.* 118 (11): 4775–4777.

31 Bäppler, S.A., Plasser, F., Wormit, M., and Dreuw, A. (2014). Exciton analysis of many-body wave functions: bridging the gap between the quasiparticle and molecular orbital pictures. *Phys. Rev. A* 90 (5): 052521.

32 Head-Gordon, M., Maurice, D., and Oumi, M. (1995). A perturbative correction to restricted open shell configuration interaction with single substitutions for excited states of radicals. *Chem. Phys. Lett.* 246 (1-2): 114–121.

33 Norman, P. and Dreuw, A. (2018). Simulating X-ray spectroscopies and calculating core-excited states of molecules. *Chem. Rev.* 118 (15): 7208–7248.

34 Davidson, E.R. (1975). The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices. *J. Comput. Phys.* 17 (1): 87–94.

35 Lanczos, C. (1950). An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *J. Res. Natl. Bur. Stand.* 45 (4): 255.

36 Pulay, P. (1980). Convergence acceleration of iterative sequences. the case of scf iteration. *Chem. Phys. Lett.* 73 (2): 393–398.

37 Crawford, T. and Schaefer, H. III, (2000). An introduction to coupled cluster theory for computational chemists. *Rev. Comput. Chem.* 14.

38 Plasser, F. and Dreuw, A. (2015). High-level ab initio computations of the absorption spectra of organic iridium complexes. *Chem. A Eur. J.* 119 (6): 1023–1036.

39 Schreiber, M., Silva-Junior, M.R., Sauer, S.P.A., and Thiel, W. (2008). Benchmarks for electronically excited states: CASPT2, CC2, CCSD, and CC3. *J. Chem. Phys.* 128 (13): 134110.

40 Silva-Junior, M.R., Schreiber, M., Sauer, S.P.A., and Thiel, W. (2008). Benchmarks for electronically excited states: time-dependent density functional theory and density functional theory based multireference configuration interaction. *J. Chem. Phys.* 129 (10): 104103.

41 Loos, P.F., Scemama, A., Blondel, A. et al. (2018). A mountaineering strategy to excited states: highly accurate reference energies and benchmarks. *J. Chem. Theory Comput.* 14 (8): 4360–4379.

42 Trofimov, A.B. and Schirmer, J. (1995). An efficient polarization propagator approach to valence electron excitation spectra. *J. Phys. B Atomic Mol. Phys.* 28 (12): 2299–2324.

43 Starcke, J.H., Wormit, M., Schirmer, J., and Dreuw, A. (2006). How much double excitation character do the lowest excited states of linear polyenes have? *Chem. Phys.* 329 (1-3): 39–49.

44 Starcke, J.H., Wormit, M., and Dreuw, A. (2009). Unrestricted algebraic diagrammatic construction scheme of second order for the calculation of excited states of medium-sized and large molecules. *J. Chem. Phys.* 130 (2): 024104.

45 Sinha, D., Mukhopadhyay, S., Chaudhuri, R., and Mukherjee, D. (1989). The eigenvalue-independent partitioning technique in Fock space: an alternative route to open-shell coupled-cluster theory for incomplete model spaces. *Chem. Phys. Lett.* 154 (6): 544–549.

46 Stanton, J.F. and Gauss, J. (1994). Analytic energy derivatives for ionized states described by the equation-of-motion coupled cluster method. *J. Chem. Phys.* 101 (10): 8938–8944.

47 Krylov, A.I. (2001). Size-consistent wave functions for bond-breaking: the equation-of-motion spin-flip model. *Chem. Phys. Lett.* 338 (4-6): 375–384.

48 Krylov, A.I. (2006). Spin-flip equation-of-motion coupled-cluster electronic structure method for a description of excited states, bond breaking, diradicals, and triradicals. *Acc. Chem. Res.* 39 (2): 83–91.

49 Lefrancois, D., Wormit, M., and Dreuw, A. (2015). Adapting algebraic diagrammatic construction schemes for the polarization propagator to problems with multi-reference electronic ground states exploiting the spin-flip ansatz. *J. Chem. Phys.* 143 (12): 124107.

50 Lefrancois, D., Tuna, D., Martínez, T.J., and Dreuw, A. (2017). The spin-flip variant of the algebraic-diagrammatic construction yields the correct topology of S 1 /S 0 conical intersections. *J. Chem. Theory Comput.* 13 (9): 4436–4441.

51 Lefrancois, D., Rehn, D.R., and Dreuw, A. (2016). Accurate adiabatic singlet-triplet gaps in atoms and molecules employing the third-order spin-flip algebraic diagrammatic construction scheme for the polarization propagator. *J. Chem. Phys.* 145 (8): 084102.

52 Winter, N.O.C. and Hättig, C. (2011). Scaled opposite-spin CC2 for ground and excited states with fourth order scaling computational costs. *J. Chem. Phys.* 134 (18): 184101.

53 Krauter, C.M., Pernpointner, M., and Dreuw, A. (2013). Application of the scaled-opposite-spin approximation to algebraic diagrammatic construction schemes of second order. *J. Chem. Phys.* 138 (4): 044107.

54 Cederbaum, L.S., Domcke, W., and Schirmer, J. (1980). Many-body theory of core holes. *Phys. Rev. A* 22 (1): 206–222.

55 Wenzel, J., Wormit, M., and Dreuw, A. (2014). Calculating core-level excitations and x-ray absorption spectra of medium-sized closed-shell molecules with the algebraic-diagrammatic construction scheme for the polarization propagator. *J. Comput. Chem.* 35 (26): 1900–1915.

56 Wenzel, J., Wormit, M., and Dreuw, A. (2014). Calculating X-ray absorption spectra of open-shell molecules with the unrestricted algebraic-diagrammatic construction scheme for the polarization propagator. *J. Chem. Theory Comput.* 10 (10): 4583–4598.

57 Wenzel, J., Holzer, A., Wormit, M., and Dreuw, A. (2015). Analysis and comparison of CVS-ADC approaches up to third order for the calculation of core-excited states. *J. Chem. Phys.* 142 (21): 214104.

58 Wenzel, J. and Dreuw, A. (2016). Physical properties, exciton analysis, and visualization of core-excited states: an intermediate state representation approach. *J. Chem. Theory Comput.* 12 (3): 1314–1330.

59 Mai, S., Plasser, F., Pabst, M. et al. (2017). Surface hopping dynamics including intersystem crossing using the algebraic diagrammatic construction method. *J. Chem. Phys.* 147 (18): 184109.

60 Mewes, J.M., You, Z.Q., Wormit, M. et al. (2015). Experimental benchmark data and systematic evaluation of two a posteriori, polarizable-continuum corrections for vertical excitation energies in solution. *Chem. A Eur. J.* 119 (21): 5446–5464.

61 Mewes, J.M., Herbert, J.M., and Dreuw, A. (2017). On the accuracy of the general, state-specific polarizable-continuum model for the description of correlated ground- and excited states in solution. *Phys. Chem. Chem. Phys.* 19 (2): 1644–1654.

62 Karbalaei Khani, S., Marefat Khah, A., and Hättig, C. (2018). COSMO-RI-ADC(2) excitation energies and excited state gradients. *Phys. Chem. Chem. Phys.* 20 (24): 16 354–16 363.

63 Scheurer, M., Herbst, M.F., Reinholdt, P. et al. (2018). Polarizable embedding combined with the algebraic diagrammatic construction: tackling excited states in biomolecular systems. *J. Chem. Theory Comput.* 14 (9): 4870–4883.

64 Sen, R., Dreuw, A., and Faraji, S. (2018). Algebraic diagrammatic construction for the polarization propagator in combination with effective fragment potentials. *Phys. Chem. Chem. Phys.* 21: 3683–3694.

65 Prager, S., Zech, A., Aquilante, F. et al. (2016). First time combination of frozen density embedding theory with the algebraic diagrammatic construction scheme for the polarization propagator of second order. *J. Chem. Phys.* 144 (20): 204103.

66 Prager, S., Zech, A., Wesolowski, T.A., and Dreuw, A. (2017). Implementation and application of the frozen density embedding theory with the algebraic diagrammatic construction scheme for the polarization propagator up to third order. *J. Chem. Theory Comput.* 13 (10): 4711–4725.

6

Foundation of Multi-Configurational Quantum Chemistry

Giovanni Li Manni, Kai Guther, Dongxia Ma, and Werner Dobrautz

Max Planck Institute for Solid State Research, Heisenbergstrasse 1, 70569 Stuttgart, Germany

Abstract

After introducing the fundamental goals—solving the Schrödinger equation—and the associated problems of quantum chemistry, we describe the basics of multiconfigurational approaches to solve the latter. As an exact—or full configuration interaction (FCI)—solution, even in a finite basis set, comes with an exponential scaling cost, the importance of an efficient representation in either a Slater determinant or configuration state function basis is discussed. With the help of such an efficient representation it is possible to apply iterative techniques, like the Davidson method, to obtain the exact solution of the most important low-lying eigenstates of the Hamiltonian, describing a quantum chemical system. As the exponential scaling still restricts these direct approaches to rather modest system sizes, we discuss in depth the multi-configurational extension of the self-consistent field method (MCSCF), which captures the static correlation of a problem and serves as a starting point for many more elaborate techniques. In addition, we present the complete active space approach—and the generalized and restricted extensions thereof—, which allows an intuitive construction of the chemically important reference space and enables a much more compact description of the important degrees of freedom of a problem at hand. We explain the state-specific and state-averaged approaches to obtain excited states within the MCSCF method and conclude this chapter by presenting stochastic Monte-Carlo approaches to solve the FCI problem for unprecedented active space sizes.

When describing chemical systems at the quantum mechanical level, electronic properties are derived from an N-electron wave function, Ψ, a complex-valued function depending on the position and spin of all electrons, that fully captures the stationary state of the system and its properties, such as the energy, E. The electronic wave function, Ψ, is determined by solving the non-relativistic time-independent Schrödinger equation

$$\hat{H}\Psi = E\Psi, \tag{6.1}$$

where $\hat{H}$ is the Hamiltonian operator describing the system. The properties of the system are obtained from expectation values of suitable operators with respect to the wave function, Ψ. Solving the Schrödinger equation exactly is extremely difficult, and the exact solution is available only for a limited number of special cases.

Even though the stationary Schrödinger equation has been formulated in real-space, and the N-electron wave function, Ψ, is a continuous function of $3N$ space variables and spin,

Quantum Chemistry and Dynamics of Excited States: Methods and Applications,
First Edition. Edited by Leticia González and Roland Lindh.

many methods rely on the discretization of the wave function, with the atomic orbitals (AOs) representing the one-particle basis, and the *Slater determinants* (SDs) the N-electron basis. In modern methods, the discretization of the wave function starts from the selection of a finite list of AOs, and in doing so an unavoidable basis set truncation error is introduced. From the AO basis, molecular orbitals are obtained and the N-electron wave function $|\Psi\rangle$ is built, using one or more SDs. The discretized wave function takes the following simple form

$$|\Psi\rangle = \sum_i C_i |i\rangle \,, \tag{6.2}$$

with $|i\rangle = |\chi_{i_1\sigma_1} \cdots \chi_{i_n\sigma_n}\rangle$ being the SDs used as N-electron basis, and C_i the corresponding coefficients. For convenience, the short notation $|i\rangle$ is used from now to denote SDs instead of the full expansion. The vector $\mathbf{C} = \{C_i\}$ is referred to as the *configuration interaction* (CI) vector and obtaining the CI vector of the ground state is equivalent to solving the Schrödinger equation within the given N-electron basis for the lowest eigenvalue.

Replacing the continuous function, Ψ, with the finite N-electron expansion of Eq. (6.2), the Schrödinger equation can be reformulated into a linear algebra problem. Starting from Eq. (6.1) and multiplying with $\langle j|$ from the left, one obtains

$$\sum_i \langle j|\hat{H}|i\rangle \; C_i = E \sum_i \langle j|i\rangle \; C_i \tag{6.3}$$

and by using an orthonormal basis set, with $\langle i|j\rangle = \delta_{ij}$, Eq. (6.3) will be simplified as

$$\sum_i \langle j|\hat{H}|i\rangle \; C_i = E \; C_j. \tag{6.4}$$

Various ways of choosing the N-electron basis (how many and which SDs) lead to different levels and flavors of quantum chemistry methods, and an accurate description of the many-body wave function can be obtained with a sensible choice of the underlying basis set, while having the striking advantage that the Schrödinger equation, previously a partial differential equation, is now a finite-dimensional eigenvalue problem.

Possible approaches to solve the eigenvalue problem of Eq. (6.4) will be discussed in Section 6.4.

Expanding the many-electron wave function on the basis of SDs has the big advantage of making the evaluation of non-vanishing $\langle j|\hat{H}|i\rangle$ terms extremely simple (see Section 6.2.1). However, Slater determinants are in general not eigenfunctions of the total spin operator, $\hat{\mathbf{S}}^2$, and in many cases solutions of the CI eigenvalue problems are obtained which are not eigenvectors of $\hat{\mathbf{S}}^2$ (*spin contamination*), making any characterization of the molecular system in terms of spin impossible. To avoid this limitation the discretization of wave functions can be carried out on the basis of spin-adapted functions, also referred to as configuration state functions (CSFs). CSFs can be constructed as a particular linear combination of SDs, that preserves the total spin. Construction and coupling of CSFs will be discussed in greater detail in Section 6.3.

Only the combination of an increasingly large one-electron basis, up to the complete basis set (CBS) limit, and the use of more elaborate methods approaching the *full configuration interaction* (FCI) limit, yields the exact solution of the given non-relativistic Schrödinger equation, as depicted in Figure 6.1. However, the number of the N-electron functions, $|i\rangle$ for a given one-electron basis, quickly becomes prohibitively large, for chemical systems of practical interest.

A configurational space of all possible SDs arising from the distribution of all the electrons in the available MO space is known as *full configuration interaction* (FCI). FCI is the theoretical limit for a given one-electron basis set. Since FCI accounts for all possible degrees of freedom, it is invariant with respect to any orbital rotation.

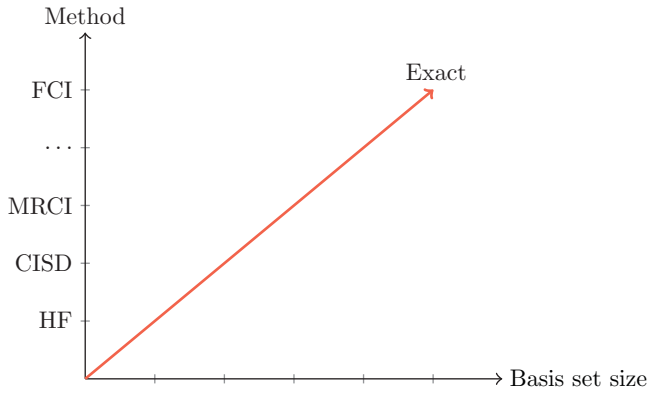


Figure 6.1 The exact solution of the Schrödinger equation of a system is approached by increasing the basis set and using of more accurate methods at the same time.

In the simple Hartree–Fock (HF) method, for the selected one-electron basis set, MOs are variationally optimized with respect to the field generated by a single SD, the HF determinant. We refer to it as a *single-configurational* approach. The Hartree–Fock wave function is invariant with respect to occupied–occupied and virtual–virtual orbital rotations. However, any occupied–virtual orbital mixing will displace the HF wave function from its variational minimum. One may enlarge the determinantal expansion and optimize the CI expansion coefficients, C_i, while keeping the HF orbitals (or any other suitable choice of MOs) fixed. This approach is at the core of the configuration interaction (CI) methods, discussed in some details in Section 6.4.1. Using a space that consists of all single and double excitations from one reference configuration, usually the HF determinant, is denoted as configuration interaction singles and doubles (CISD). This method has the same invariance with respect to orbital rotations as HF. A determinantal expansion generated by exciting a selected number of *active electrons*, N, in a selected subset of *active orbitals*, n, around the *Fermi level* (frontier orbitals) in all possible ways, compatible with space and spin symmetry constraints, leads to the *complete active space* configuration interaction, CASCI(N,n), wave function. More options exist and some of them will be discussed in Section 6.6. In the *multi-configurational self-consistent field (MCSCF) methods*, both expansion coefficients and MOs are simultaneously variationally optimized. As opposed to the Hartree–Fock method, in MCSCF the MOs are optimized under the averaged field generated by a multi-configurational wave function. The MCSCF wave function parametrization and its optimization will be discussed in great detail in Section 6.5. A commonly used MCSCF method is the *complete active space self-consistent field* (CASSCF) approach, where the CI expansion is obtained as the full-CI in a suitable active space. The main drawback of the CASSCF method is that the CAS wave function grows exponentially with the size of the active space, and for practical simulations, the current computational limit on the size of the active space is of about 18 electrons and 18 orbitals, CAS(18,18). One approach that is used to partially circumvent the exponential scaling is to use truncated wave functions. Various forms of truncations of the CAS wave function have been successfully applied, such as the restricted active space (RAS) and the generalized active space (GAS) wave functions leading to the RASSCF and the GASSCF approaches. RAS and GAS wave functions will be discussed in Section 6.6, with a focus on how they are constructed and applied in various branches of quantum chemistry.

In recent years, alternative methods to tackle the exponential scaling in CASSCF calculations have been developed, with the Density Matrix Renormalization Group (DMRG-SCF) method and the Stochastic-CASSCF being promising examples. The Stochastic-CASSCF method is based on the FCI Quantum Monte Carlo approach (FCIQMC) as the CI eigensolver. The DMRG method is discussed in Chapter (7) while FCIQMC and Stochastic-CASSCF will be discussed in Section 6.8.

6.1 Scaling Problem in FCI, CAS and RAS Wave Functions

The major drawback of FCI and CAS methods is the exponential scaling behavior with the number of correlated orbitals and electrons, such that, even for small active spaces, the size of the considered Hilbert space can be prohibitively large.

The total number of Slater determinants, N_{SD}, that can be generated by considering all possible distributions of N electrons in n orbitals with a total spin projection eigenvalue M_s and without imposing space symmetry constraints is given by

$$N_{SD}(N, n, M_s) = \binom{n}{\frac{N}{2} + M_s} \binom{n}{\frac{N}{2} - M_s}. \tag{6.5}$$

When expressed in terms of α and β electrons, this equation can be written as

$$N_{SD}(n, N_\alpha, N_\beta) = \binom{n}{N_\alpha} \binom{n}{N_\beta} \tag{6.6}$$

as $N = N_\alpha + N_\beta$ and $M_s = (N_\alpha - N_{\beta)})/2$.

The number of Slater determinants and the memory required to store the corresponding CI vector for a series of complete active space sizes, CAS(N,n), with $N = n$ and $S = M_s = 0$ is listed in Table 6.1 showing that there is a tight technical limit on how many electrons and orbitals can be explicitly correlated in multi-configurational methods. Even though additional symmetry constraints, such as point group symmetry, can reduce the size of the wave function, this is typically only by one order of magnitude, such that the accessible system sizes cannot be substantially increased.

Slater determinants are not eigenfunctions of the total spin operator, $\hat{S}^2$, and Eq. (6.6) represents an upper bound to the number of spin eigenfunctions. The non-relativistic Hamiltonian $\hat{H}$ from Chapter 1 (Eq. (1.18)) is spin independent, hence, all states with a given total spin quantum number S are degenerate, regardless of their M_S eigenvalue. It is therefore sufficient to consider only one

Table 6.1 Number of SDs and memory requirement for different active space choices.

(*N*,*n*)	N_{SDs}	Memory per vector
(8,8)	4900	~38.3 kB
(12,12)	853776	~6.51 MB
(16,16)	165636900	~1.23 GB
(20,20)	34134779536	~254.32 GB
(24,24)	7312459672336	~53.21 TB

Number of slater determinants N_{SDs} and memory requirement for storing a single CI vector for a problem of N electrons in n orbitals with $N = n$ and $m_s = 0$ and assuming a size of 8 bytes per entry of the CI vector.

state per value of S, which reduces the size of the targeted space. The number of spin-adapted electronic configurations, referred to as *configuration state functions*, CSFs, with a specific total spin quantum number $S = M_s$, are obtained by removing the determinants with higher spin from Eq. (6.6)

$$\begin{aligned} N_{CSF}(N,n,S) &= N_{SD}(n,N_\alpha,N_\beta) - N_{SD}(n,N_\alpha+1,N_\beta-1) \\ &= \binom{n}{N_\alpha}\binom{n}{N_\beta} - \binom{n}{N_\alpha+1}\binom{n}{N_\beta-1} \end{aligned} \tag{6.7}$$

The Weyl–Paldus dimension formula can also be used to count the total number of spin adapted functions

$$N_{CSF}(N,n,S) = \frac{2S+1}{n+1}\binom{n+1}{\frac{N}{2}-S}\binom{n+1}{n-\frac{N}{2}-S}. \tag{6.8}$$

Note that Eq. (6.8) and Eq. (6.7) are equivalent.

The number of determinants and configuration state functions for a series of complete active space sizes, CAS(N,n), with $N = n$ and $S = M_s = 0$ is also reported in Figure 6.2. The gray area represents the zone where the calculations require significant computational resources and dedicated large-memory machines. Beyond the gray area, calculations become technically impossible nowadays. If Stirling's approximation is used, $n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$, the equations for the total number of SDs and CSFs can be approximated (with $N = n$) by

$$N_{SD} \cong \frac{2}{\pi n}4^n \tag{6.9}$$

$$N_{CSF} \cong \left(1 - \left(\frac{n}{n+2}\right)^2\right)\frac{2}{\pi n}4^n. \tag{6.10}$$

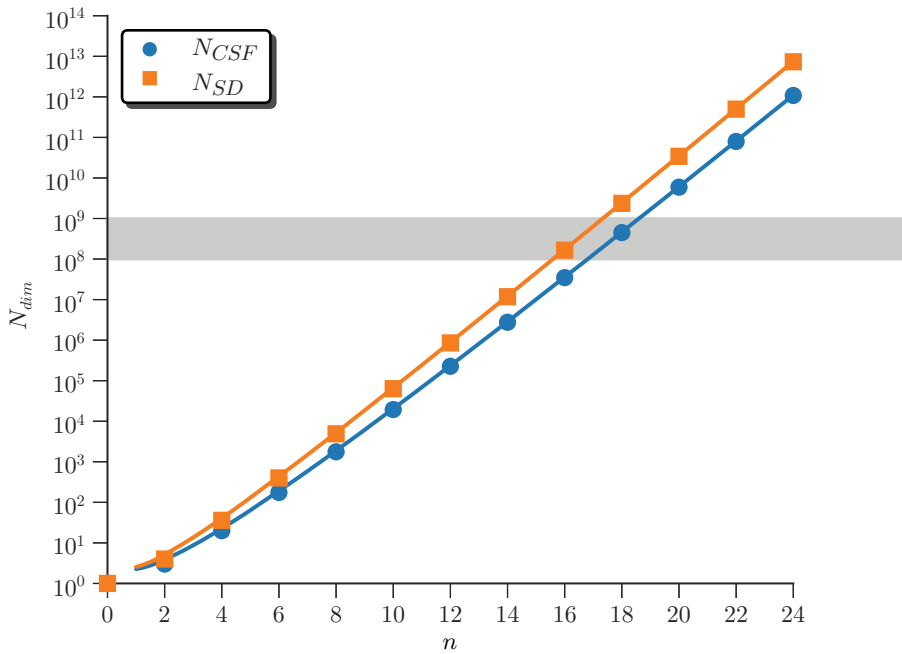


Figure 6.2 Number of determinants (orange squares) and CSFs (blue circles) for a complete active space of N electrons in n orbitals, where $N = n$ and $M_s = 0$.

Eq. (6.9) and Eq. (6.10) clearly show that the number of SDs and CSFs grows exponentially with the number of orbitals in the active space. The growth of SDs and CSFs approximated by Stirling's formula is represented by solid lines in Figure 6.2. The exponential scaling can also be demonstrated via a more qualitative argument, by considering that the number of ways one can fill n orbitals with electrons is equal to 4^n, as each orbital can be either doubly occupied, singly occupied with one α or one β electron, or unoccupied, independent of the occupancy of the other orbitals. This generates all possible configurations of any electron number in n orbitals, but the scaling behavior is qualitatively the same as for a fixed number of electrons. The use of the term "factorial scaling" referring to the binomial coefficient is not recommended as this would imply a much faster growth than the 4^n growth demonstrated above.

6.2 Factorization and Coupling of Slater Determinants

The exponential scaling of FCI and CASCI wave functions quickly makes dense matrix operations prohibitively expensive. Iterative methods (see Section 6.4.1), such as the Davidson method [1, 2], have been recommended to avoid full diagonalization of large Hamiltonian matrices. These methods require storage of the σ-*vector*, $\sigma = HC$, instead of the full Hamiltonian matrix. In order for the contraction to the σ-vector to be practical its evaluation must be done efficiently. Methods for an efficient evaluation of the sigma vector have been reported by Siegbahn [3], Knowles and Handy [4].

In this section we will discuss Handy's technique [5] of separating Slater determinants in α-strings and β-strings, which represents a major milestone in determinant based CI techniques. This procedure is used to generate determinants in a well defined order and to evaluate density matrix or sigma-vector contributions in a computationally advantageous manner. The strength of this approach will be made more clear in Section 6.4.2, when the Direct-CI algorithm will be discussed [6, 7].

According to Handy's approach a Slater determinant is factorized as follows

$$|\alpha(I_\alpha)\beta(I_\beta)\rangle = \hat{\alpha}(I_\alpha)\hat{\beta}(I_\beta)|vac\rangle. \tag{6.11}$$

The α-string, $\hat{\alpha}(I_\alpha)$, and the β-string, $\hat{\beta}(I_\beta)$, are products of n_α and n_β creation operators for α-spin-orbitals and β-spin-orbitals, respectively. The dimension of each string is a constant and it is defined by the total number of electrons, N, and the spin projection, M_s (z component of the total spin) for the given system

$$n_\alpha = (N + 2M_s)/2 \tag{6.12}$$

$$n_\beta = (N - 2M_s)/2. \tag{6.13}$$

The number of α- and β-strings is obtained by the binomial coefficients

$$N^\alpha_{strings} = \begin{pmatrix} n \\ N_\alpha \end{pmatrix} \tag{6.14}$$

$$N^\beta_{strings} = \begin{pmatrix} n \\ N_\beta \end{pmatrix} \tag{6.15}$$

with N_α (or N_β) electrons distributed in n orbitals. Their product gives the total number of Slater determinants

$$N_{det} = \begin{pmatrix} n \\ N_\alpha \end{pmatrix} \begin{pmatrix} n \\ N_\beta \end{pmatrix} \tag{6.16}$$

as seen already in the previous section. The factorization of strings allows Slater determinants to be read, processed and stored in a matrix form. This approach makes the evaluation of operators acting only on α (or β) electrons highly efficient. Identically, CI vectors and sigma vectors can be vectorized and stored in matrix form.

A graphical representation of strings is used to order the strings. String ordering is advantageous as the position of a string in a string-list provides information on the distribution of the electrons in orbitals. We consider the case of N α-electrons in n orbitals. Each string can be represented as a path in an $n \times N$ graph, obtained by drawing arcs between vertices (k, m), where k is the orbital index and m the number of electrons in the orbitals up to orbital k (see Figure 6.3). All paths start at (0,0) and end at (n, N). A vertical arc on the path from vertex (k, m) to $(k + 1, m)$ means that orbital $(k + 1)$ is unoccupied. A diagonal arc from vertex (k, m) to $(k + 1, m + 1)$ means that orbital $(k + 1)$ is occupied. For example, the α-string of 3 electrons in 5 orbitals, $\hat{a}^\dagger_{1\alpha}\hat{a}^\dagger_{3\alpha}\hat{a}^\dagger_{5\alpha}|vac\rangle$ can be written as a vector

$$\hat{a}^\dagger_{\mathbf{1}\alpha}\hat{a}^\dagger_{\mathbf{3}\alpha}\hat{a}^\dagger_{\mathbf{5}\alpha}|\mathbf{vac}\rangle = \begin{pmatrix} k & m \\ 1 & 1 \\ 2 & 1 \\ 3 & 2 \\ 4 & 2 \\ 5 & 3 \end{pmatrix} \tag{6.17}$$

and graphically represented by the path in Figure 6.3.

Strings are ordered in *reverse lexical order*, that is, string X comes before string Y if, in the *last occupation* where they differ, string X has a lower orbital number. For instance, in distributing 3 α-electrons into 5 orbitals, α-string $a^\dagger_{1\alpha}a^\dagger_{2\alpha}a^\dagger_{4\alpha}|vac\rangle$ comes before α-string $a^\dagger_{1\alpha}a^\dagger_{3\alpha}a^\dagger_{5\alpha}|vac\rangle$. In order to obtain the reverse lexical ordering from the graphical representation of strings, a *vertex weight* $W_{k,m}$ is associated to each *allowed* vertex (a vertex is called allowed if it is visited by at least one path) equal to the number of different paths from (0,0) to (k, m). Since all such paths must come from

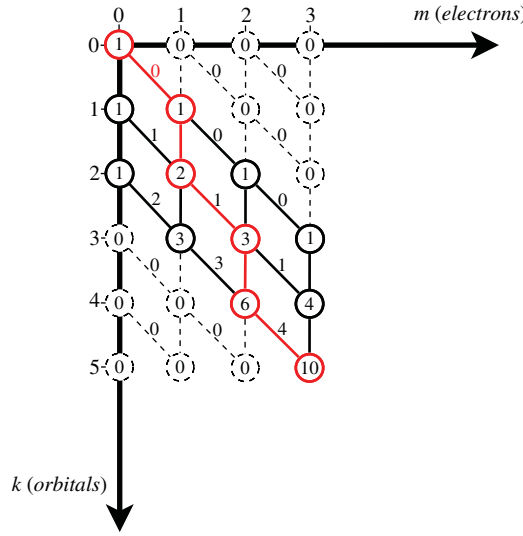


Figure 6.3 Path describing string $a^\dagger_{1\alpha}a^\dagger_{3\alpha}a^\dagger_{5\alpha}|vac\rangle$ (in red). The path contains three diagonal arcs (occupied orbitals) and two vertical arcs (unoccupied orbitals).

either $(k-1, m-1)$ or $(k-1, m)$, the following recurrence relation between the vertex weights (Eq. (6.18)) is derived

$$W_{k,m} = W_{k-1,m} + W_{k-1,m-1} \tag{6.18}$$

This equation clearly shows that the weight of each vertex is given by the sum of the weights of the vertex right above and top-left. The head vertex (0,0) has a weight equal to one ($W_{0,0} = 1$). Weights of *forbidden* vertices are set to *zero*. All other vertex weights are computed from Eq. (6.18). We also introduce the *arc weights* using the following recursive relation

$$Y_{k+1,m+1} = W_{k+1,m+1} - W_{k,m} = W_{k,m+1} \tag{6.19}$$

which is equal to the vertex weight located on the top-right side of the arc. Arc weights of zero are attributed to vertical arcs. A *path weight* is then obtained as a sum of arc weights along the considered path. For the example given in Figure 6.3, we find that the path weight is $I_\alpha = 5$. Path weights represent the indexing number of reverse lexically ordered determinants. With path weights in hand, any list of strings can be written and sorted according to the reverse lexical order. α- and β-strings with lower path weight occur earlier in the ordered string list. We refer to the literature for further details on the topic [7, 8].

6.2.1 Slater Condon Rules

An efficient way to evaluate Hamiltonian matrix elements

$$H_{ij} = \langle D_i|\hat{H}|D_j\rangle \tag{6.20}$$

between Slater determinants is essential in quantum chemistry applications. As the electronic Hamiltonian contains at most two-body interactions, any matrix element between two determinants $|D_i\rangle$ and $|D_j\rangle$ that differ by more than four spin-orbitals vanishes.

First, suppose $|D_i\rangle$ and $|D_j\rangle$ differ by exactly four spin-orbitals, that is, there are two spin-orbitals R,S[1] that are occupied in $|D_i\rangle$ and not in $|D_j\rangle$ and two orbitals P,Q that are occupied in $|D_j\rangle$ and not in $|D_i\rangle$, while all other spin-orbitals have identical occupation number in $|D_i\rangle$ and $|D_j\rangle$. We refer to $|D_j\rangle$ as a double excitation of $|D_i\rangle$, and it can then be written as

$$|D_j\rangle = \hat{a}_P^\dagger \hat{a}_Q^\dagger \hat{a}_R \hat{a}_S |D_i\rangle \ . \tag{6.21}$$

Inserting the expression of the Hamiltonian operator from Chapter 1, Eq. (1.18), into the element $\langle D_j|\hat{H}|D_i\rangle$ and considering that all terms not involving the double excitation $\hat{a}_P^\dagger \hat{a}_Q^\dagger \hat{a}_R \hat{a}_S$ vanish yields

$$\begin{aligned} H_{ij} &= \frac{1}{2}((pr|qs)\langle D_j|\hat{E}_{pr}\hat{E}_{qs}|D_i\rangle + (qs|pr)\langle D_j|\hat{E}_{qs}\hat{E}_{pr}|D_i\rangle \\ &\quad + (ps|qr)\langle D_j|\hat{E}_{ps}\hat{E}_{qr}|D_i\rangle + (qr|ps)\langle D_j|\hat{E}_{qr}\hat{E}_{ps}|D_i\rangle) \\ &= (pr|qs) - (ps|qr) \end{aligned} \tag{6.22}$$

Thus, in case of a double excitation, the matrix element only depends on the excitation operator and is independent of the remaining orbitals.

For a single excitation

$$|D_j\rangle = \hat{a}_P^\dagger \hat{a}_R |D_i\rangle \ , \tag{6.23}$$

1 Capital letters indicate the combined spatial- and spin-coordinate of a spin-orbital.

the following expression is obtained

$$\begin{aligned} H_{ij} &= h_{pr}\langle D_j|\hat{E}_{pr}|D_i\rangle + \frac{1}{2}\sum_k((pr|kk)\langle D_j|\hat{E}_{pr}\hat{E}_{kk}|D_i\rangle + (kk|pr)\langle D_j|\hat{E}_{kk}\hat{E}_{pr}|D_i\rangle \\ &\quad + (pk|kr)\langle D_j|\hat{E}_{pk}\hat{E}_{kr} - \hat{E}_{pr}|D_i\rangle + (kr|pk)\langle D_j|\hat{E}_{kr}\hat{E}_{pk} - \hat{E}_{pr}|D_i\rangle) \\ &= h_{pr} + \sum_k[(pr|kk) - (pk|kr)]\langle D_i|\hat{n}_k|D_i\rangle\,, \end{aligned} \tag{6.24}$$

where $\hat{n}_k = \hat{a}_k^\dagger\hat{a}_k$ is the *occupation number operator* of spatial-orbital k, which restricts the sum to occupied orbitals. That is, in contrast to the double excitations, the matrix element now depends on which orbitals are occupied.

For diagonal matrix elements H_{ii} no contribution arises from excitation operators that change the occupancy of the spin-orbitals, and

$$\begin{aligned} H_{ii} &= \sum_k h_{kk}\langle D_i|\hat{E}_{kk}|D_i\rangle + \frac{1}{2}\sum_{kl}((kk|ll)\langle D_i|\hat{E}_{kk}\hat{E}_{ll}|D_i\rangle \\ &\quad + (ll|kk)\langle D_i|\hat{E}_{ll}\hat{E}_{kk}|D_i\rangle + 2(kl|lk)\langle D_i|\hat{E}_{kl}\hat{E}_{kl} - \hat{E}_{kl}|D_i\rangle) \\ &= \sum_k h_{kk}\langle D_i|\hat{n}_k|D_i\rangle + \sum_{kl}[(kk|ll) - (kl|lk)]\langle D_i|\hat{n}_k\hat{n}_l|D_i\rangle \end{aligned} \tag{6.25}$$

Again, the expectation values of the occupation number operators restrict the sums to occupied spatial-orbitals.

Equations (6.22–6.25) are referred to as the Slater–Condon rules and reduce the evaluation of the Hamiltonian matrix elements of an N-electron problem to the computation of one- and two-body integrals. The computational cost of evaluating a matrix element only scales linearly with the number of occupied orbitals and it is independent of the number of virtual orbitals of the involved SDs.

As a final remark, the expression of $|D_j\rangle$ as an excitation of the form (6.21) or (6.23) is essential for the applicability of the Slater–Condon rules. However, the order of the operators in Eq. (6.21) is not fixed and a different ordering of the fermionic creation/annihilation operators, i.e., $a_P^\dagger a_Q^\dagger a_S a_R$, can introduce a change of sign.

6.3 Configuration State Functions

Slater determinants are eigenfunctions of the z-component, $\hat{S}_z$, of the spin-vector operator $\hat{\mathbf{S}}$, with associated eigenvalue m_s. However, they are not eigenfunctions of the total spin operator $\hat{\mathbf{S}}^2$, as can be demonstrated by a simple two electrons in two orbitals example. We denote the two orbitals as A and B. A total of six Slater determinants can be built

$$|A(\alpha)A(\beta)\rangle,\ |B(\alpha)B(\beta)\rangle,\ |A(\alpha)B(\beta)\rangle,\ |A(\beta)B(\alpha)\rangle,\ |A(\alpha)B(\alpha)\rangle,\ |A(\beta)B(\beta)\rangle \tag{6.26}$$

with the other combinations of single-particle basis functions, such as $|A(\alpha)A(\alpha)\rangle$, being forbidden by the Pauli exclusion principle. All six determinants are eigenfunctions of $\hat{S}_z$, but only four are eigenfunctions of $\hat{\mathbf{S}}^2$, namely $|A(\alpha)A(\beta)\rangle$, $|B(\alpha)B(\beta)\rangle$ (singlet, $S = 0$, $m_s = 0$) and $|A(\alpha)B(\alpha)\rangle$, $|A(\beta)B(\beta)\rangle$ (triplet, $S = 1$, $m_s = \pm 1$). Determinants $|A(\alpha)B(\beta)\rangle$ and $|A(\beta)B(\alpha)\rangle$ are not eigenfunctions of $\hat{\mathbf{S}}^2$. Determinants containing only doubly occupied orbitals, such as $|A(\alpha)A(\beta)\rangle$ and $|B(\alpha)B(\beta)\rangle$ are called *closed-shell* determinants and are always pure singlet spin functions. Determinants containing singly occupied orbitals, such as $|A(\alpha)B(\beta)\rangle$, $|A(\beta)B(\alpha)\rangle$, $|A(\alpha)B(\alpha)\rangle$, $|A(\beta)B(\beta)\rangle$ are referred to as *open-shell* determinants. Open-shell determinants are spin eigenfunctions only when all singly occupied orbitals have parallel spin, such as in the $|A(\alpha)B(\alpha)\rangle$ case.

In all other cases open-shell determinants are not pure spin eigenfunctions. However, open-shell determinants with non-parallel spins can be utilized as bases in the diagonalization of

$$\hat{\mathbf{S}}^2 = (\hat{\mathbf{S}}_A + \hat{\mathbf{S}}_B)^2, \tag{6.27}$$

where $\hat{\mathbf{S}}_A$ and $\hat{\mathbf{S}}_B$ are the total spin operators of orbital A and B, to obtain spin-eigenfunctions. In the previous example, the two-dimensional matrix arising from the combination of $|A(\alpha)B(\beta)\rangle$ and $|A(\beta)B(\alpha)\rangle$ via the total spin operator leads to the two spin-adapted functions

$$S = 0: \quad \frac{1}{\sqrt{2}}(|A(\alpha)B(\beta)\rangle - |A(\beta)B(\alpha)\rangle), \tag{6.28}$$

$$S = 1: \quad \frac{1}{\sqrt{2}}(|A(\alpha)B(\beta)\rangle + |A(\beta)B(\alpha)\rangle), \tag{6.29}$$

The first eigenfunction, Eq. (6.28), is a singlet state ($S = 0$, $m_s = 0$). The second eigenfunction, Eq. (6.29), is a triplet state, ($S = 1$, $m_s = 0$). These spin adapted eigenfunctions are called *configuration state functions* (CSFs).

CSFs for more electrons can be created analogously, by diagonalizing $\hat{\mathbf{S}}^2$ in the subspace spanned by the "complete" set of open-shell Slater determinants with non-parallel spins. The CSFs for N electrons can be created from those for $N-1$ orbitals by adding $\hat{\mathbf{S}}_N$ to $\hat{\mathbf{S}}_{\mathbf{N-1}}$ and diagonalizing the matrix corresponding to the new total spin, $\hat{\mathbf{S}}_N$. The form of the resulting eigenvectors of $\hat{\mathbf{S}}^2$ Eqs. (6.28, 6.29) does not depend on the choice of the single-particle basis and an orbital rotation is not sufficient to enforce spin conservation.

The diagonalization of the $\hat{\mathbf{S}}^2$ operator in a full basis of SDs is not a practical approach for a large number of electrons N and more elaborate techniques have been developed for this scope.

As CSFs are linear combinations of Slater determinants, simple rules such as the Slater–Condon rules are not available to the CSF per se, although they still apply to each of the Slater determinants of which a CSF is composed. This feature represents a major challenge for computing Hamiltonian matrix elements in a CSF basis. An elegant approach to efficiently build CSFs and couple them through the Hamiltonian operator is the graphical unitary group approach (GUGA) [9, 10] that will be discussed in the next few sections.

6.3.1 The Unitary Group Approach (UGA)

The *unitary group approach* (UGA) is an elegant and efficient method to create a spin-adapted basis (CSFs) and to calculate the Hamiltonian matrix elements in this basis. The method was pioneered by Moshinsky [11], Paldus [12] and Shavitt [9, 13], who also introduced the graphical-UGA (GUGA) for practical and fast evaluation of the Hamiltonian matrix elements.

The UGA is based on the important observation that the spin-free excitation operators Eq. (1.19) in the non-relativistic Hamiltonian Eq. (1.18) (from Chapter 1) follow the same commutation relations

$$[\hat{E}_{ij}, \hat{E}_{kl}] = \delta_{jk}\hat{E}_{il} - \delta_{il}\hat{E}_{kj}, \tag{6.30}$$

as the generators of the *unitary group* $U(n)$[2], with n being the number of spatial orbitals.

Usually in CI calculations one uses a single particle basis set of $2n$ spin-orbitals, with associated $(2n)^2$ operators $a^\dagger_{k\sigma}a_{l\sigma'}$, that correspond to the $U(2n)$ group. Due to the spin-independence of the Hamiltonian Eq. (1.18), we can use the *direct product* $U(n) \times U(2) = U(2n)$, where $U(n)$

2 The unitary group of degree n is the group of all $n \times n$ unitary matrices with the matrix multiplication as the group operation.

corresponds to the spatial orbital part and $U(2)$ to the spin. To generate a spin adapted basis we only need to treat the spatial $U(n)$ subgroup, since the direct product with $U(2)$ produces spin eigenfunctions.

6.3.1.1 Analogy between CSFs and Spherical Harmonics

The connection between CSFs and the generators of the unitary group $U(n)$ is rather complex and requires a large mathematical background. A rigorous explanation goes beyond the scope of the present chapter and will not be attempted. A qualitative analogy with spherical harmonics is here made to provide the elements involved in this connection, but it can be skipped, if desired.

The components $\hat{L}_x, \hat{L}_y$ and $\hat{L}_z$, of the orbital angular momentum operator $\hat{\mathbf{L}}$ are the generators of the group of rotations in three-dimensional space, $SO(3)$. Operators which commute with all generators of a group are called *Casimir* operators. As $[\hat{\mathbf{L}}^2, \hat{L}_i] = 0$ $(i = x, y, z)$, the $\hat{\mathbf{L}}^2$ operator is the Casimir operator of $SO(3)$. At the same time $\hat{L}_z$ is the Casimir operator of the 2D rotational group, $SO(2)$. The spherical harmonics, Y_{lm}, are simultaneous eigenfunctions of these commuting Casimir operators, $\hat{\mathbf{L}}^2$ of $SO(3)$ and $\hat{L}_z$ of $SO(2)$. Therefore they are *adapted* to the group chain $SO(3) \supset SO(2)$. The values l and m, that completely define the eigenfunctions Y_{lm}, are directly related to the eigenvalues of $\hat{\mathbf{L}}^2$ and $\hat{L}_z$, respectively. Thus they specify the irreducible representation (irrep) of $SO(3)$ and $SO(2)$ respectively. Restrictions on the possible values of m for a given l apply, namely $-l \leq m \leq l$. The action of the generators of $SO(3)$, $\hat{L}_x, \hat{L}_y$ and $\hat{L}_z$, on Y_{lm} results in linear combinations of spherical harmonics with differing m, but identical l quantum number

$$\hat{L}_x Y_{lm} = \sum_{m'} c_m Y_{lm'}. \tag{6.31}$$

Thus they form an *invariant* (l does not change) and *irreducible*[3] ($\{Y_{lm}\}$ remains orthogonal) basis under the action of the generators of $SO(3)$.

A similar approach is utilized for the generators of the unitary group $U(n)$, to obtain a particular spin-adapted basis, called the *Gel'fand-Tsetlin* (GT) basis [14–16].

The group $U(n)$ has n^2 generators, $\hat{E}_{ij}$, and a total of n Casimir operators. Similar to the Y_{lm}, which are based on the group chain $SO(3) \supset SO(2)$, the GT basis is based on the group chain

$$U(n) \supset U(n-1) \supset \cdots \supset U(2) \supset U(1), \tag{6.32}$$

where $U(1)$ possesses a one-dimensional irrep Each subgroup $U(n-1), U(n-2), \ldots, U(1)$ has $n-1, n-2, \ldots, 1$ Casimir operators, resulting in a total of $n(n+1)/2$ commuting operators. The simultaneous eigenfunctions of these operators form the GT basis, uniquely labeled by a set of $n(n+1)/2$ integers related to the eigenvalues. The focus of the next section will be the practical generation of the GT basis and their properties.

6.3.1.2 Gel'fand-Tsetlin Basis

We represent a general N-electron CSF in the following table

$$[m] = \begin{bmatrix} m_{1,n} & & m_{2,n} & \cdots & m_{n-1,n} & & m_{n,n} \\ & m_{1,n-1} & & \cdots & & m_{n-1,n-1} & \\ & \ddots & & \cdots & & \therefore & \\ & & m_{1,2} & & m_{2,2} & & \\ & & & m_{1,1} & & & \end{bmatrix} \tag{6.33}$$

referred to as the *Gel'fand pattern* [14–16].

3 There is no subset that would form an invariant basis.

The integers m_{ij}, with $i \leq j$, $j = 1, 2, \ldots, n$, n being the total number of orbitals, represent the $n(n+1)/2$ eigenvalues of the Casimir operators of the unitary group subduction chain Eq. (6.32). The j index represents the level the entry is located (row number), with $j = 1$ at the bottom level and $j = n$ the uppermost. The i entry represents the position of the m_{ij} element inside the j^{th} row. The m_{ij} entries satisfy the so called *in-between condition* [17], that is

$$m_{i,j+1} \geq m_{ij} \geq m_{i+1,j+1} \tag{6.34}$$

and as a consequence

$$m_{1n} \geq m_{2n} \ \cdots \geq m_{nn} \tag{6.35}$$

Thus, the entries in each row are sorted in decreasing order. This is analogous to the in-between condition in angular momentum eigenvalues $\{l, m_l\}$ with $-l \leq m_l \leq l$.

The n non-increasing integers of the top row of Eq. (6.33), $[m]_n = [m_{1n}, m_{2n}, \ldots, m_{nn}]$, are called the *highest weights* of the representation, and uniquely define the state to represent, just as l does for the spherical harmonics. The following $n-1$ rows define the states belonging to the chosen irrep, similar to m_l in the case of Y_{lm}.

When the Gel'fand tables are used for building electronic wave functions for an (N-electrons, n-orbitals) system, m_{ij} entries are restricted to

$$0 \leq m_{ij} \leq 2 \tag{6.36}$$

due to the Pauli exclusion principle. Two more conditions are enforced for spin and particle conserving systems

$$\sum_{i=1}^{n} m_{in} = N, \tag{6.37}$$

$$\sum_{i=1}^{n} \delta_{1,m_{in}} = 2S. \tag{6.38}$$

to ensure that the number of electron N distributed among the n space orbitals is kept constant, and that the number of singly occupied orbitals in the considered state is equal to twice the total spin, S.

Therefore the electronic Gel'fand Table 6.39

$$\begin{bmatrix}
2 & 2 & 1 & 1 & 0 & 0 & 0 & 0 \\
 & 2 & 2 & 1 & 1 & 0 & 0 & 0 \\
 & & 2 & 1 & 1 & 0 & 0 & 0 \\
 & & & 2 & 1 & 1 & 0 & 0 \\
 & & & & 2 & 1 & 0 & 0 \\
 & & & & & 1 & 1 & 0 \\
 & & & & & & 1 & 0 \\
 & & & & & & & 1
\end{bmatrix} \tag{6.39}$$

represents a system with $n = 8$ (the table contains eight rows), $N = 6$ (the sum of the top row entries is equal to six) and $S = 1$ (the sum of the singly occupied orbitals is equal to two).

As a second example, the top row of $[m]_3 = [2, 1, 0]$ specifies an electronic state with 3 electrons (the sum of the top row entries is equal to 3) in 3 spatial orbitals (3 entries in the top-row) with a total spin of $S = 1/2$ (the sum of the singly occupied orbitals is equal to 1). All CSFs belonging to this irrep are obtained by filling the Gel'fand pattern with integers m_{ij} in all possible ways, fulfilling the *in-between* condition Eq. (6.34). These are shown in Table 6.2.

Table 6.2 The 8 CSFs belonging to the irrep $[m]_3 = [2, 1, 0]$, corresponding to $N = 3$, $n = 3$ and $S = 1/2$, represented by their Gel'fand patterns.

$\begin{bmatrix} 2 & & 1 & & 0 \\ & 2 & & 1 & \\ & & 2 & & \end{bmatrix}$	$\begin{bmatrix} 2 & & 1 & & 0 \\ & 2 & & 1 & \\ & & 1 & & \end{bmatrix}$	$\begin{bmatrix} 2 & & 1 & & 0 \\ & 2 & & 0 & \\ & & 2 & & \end{bmatrix}$	$\begin{bmatrix} 2 & & 1 & & 0 \\ & 2 & & 0 & \\ & & 1 & & \end{bmatrix}$
$\begin{bmatrix} 2 & & 1 & & 0 \\ & 2 & & 0 & \\ & & 0 & & \end{bmatrix}$	$\begin{bmatrix} 2 & & 1 & & 0 \\ & 1 & & 1 & \\ & & 1 & & \end{bmatrix}$	$\begin{bmatrix} 2 & & 1 & & 0 \\ & 1 & & 0 & \\ & & 1 & & \end{bmatrix}$	$\begin{bmatrix} 2 & & 1 & & 0 \\ & 1 & & 0 & \\ & & 0 & & \end{bmatrix}$

6.3.1.3 Paldus and Weyl Tables

Considering that entries in Tables 6.39 and 6.2 assume only values 0, 1 and 2, a more compact "*three-column table*" can be utilized

$$\begin{bmatrix} & & & & & & & \\ 2 & 2 & 1 & 1 & 0 & 0 & 0 & 0 \\ & 2 & 2 & 1 & 1 & 0 & 0 & 0 \\ & & 2 & 1 & 1 & 0 & 0 & 0 \\ & & & 2 & 1 & 1 & 0 & 0 \\ & & & & 2 & 1 & 0 & 0 \\ & & & & & 1 & 1 & 0 \\ & & & & & & 1 & 0 \\ & & & & & & & 1 \end{bmatrix} \equiv \begin{bmatrix} a_i & b_i & c_i \\ 2 & 2 & 4 \\ 2 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 2 \\ 1 & 1 & 2 \\ 0 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \tag{6.40}$$

counting the number of $2's$, $1's$ and $0's$ in each row i, denoted by a_i, b_i and c_i. They are referred to as Paldus tables [12]. For each row

$$a_i + b_i + c_i = i, \quad (i = 1, \ldots, n), \tag{6.41}$$

so any two columns are sufficient to uniquely determine the state. Paldus AC tables can also be drawn as in Figure 6.4. where the number of blocks on the left and right sides of the red "*spine*" represent the entries in the AC table for each level.

The top row of any Paldus table satisfies the following properties

$$a = a_n = \frac{1}{2}N - S \tag{6.42}$$

$$b = b_n = 2S \tag{6.43}$$

$$c = c_n = n - a - b = n - \frac{1}{2}N - S \tag{6.44}$$

The Paldus table emphasizes the cumulative aspects of the coupling between electrons, with the ith row providing information on number of electrons, N_i (up to ith level) and the spin, S_i

$$N_i = 2a_i + b_i \tag{6.45}$$

$$S_i = \frac{1}{2}b_i, \tag{6.46}$$

with the restriction that for any i the intermediate value of the total spin

$$\sum_{j=1}^{i} S_j \geq 0. \tag{6.47}$$

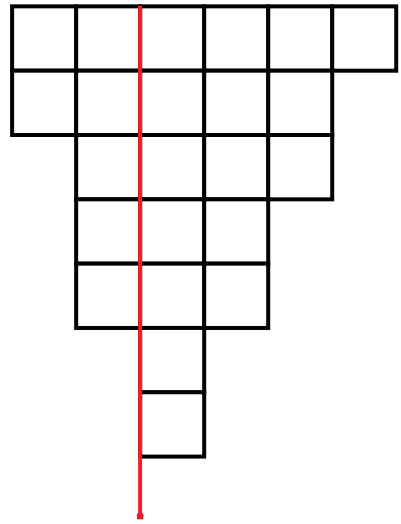


Figure 6.4 Paldus AC table representing the distribution of 6 electrons in 8 orbitals and $S = 1$, as in Table 6.40. The number of blocks on the left and right sides of the red "*spine*" represent the entries in the AC table for each level.

Paldus tables can be recast into *variation-tables* with $\Delta x_i = x_i - x_{i-1}$ $(x = a, b, c)$ as in Table 6.48.

$$\left[\begin{array}{ccc} a_i & b_i & c_i \\ \hline 2 & 2 & 4 \\ 2 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 2 \\ 1 & 1 & 2 \\ 0 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{array}\right] \rightarrow \left[\begin{array}{ccc} \Delta a_i & \Delta b_i & \Delta c_i \\ \hline 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & -1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{array}\right] \tag{6.48}$$

The Δa and Δc entries are restricted to 1 and 0 values, while Δb may assume 1, 0 and -1. By construction $\Delta a + \Delta b + \Delta c = 1$, and at each new level (from bottom to top) only one action of the following is permitted: (1) add an empty orbital, (2) add a singly occupied orbital, (3) add a doubly occupied orbital, (4) add a doubly occupied and an empty orbital and remove a singly occupied orbital. The last composite action occurs when $\Delta b = -1$. Analogously, given the row, i, of the Paldus AC table, (see Figure 6.4), at most four actions can follow for obtaining the lower row, $(i - 1)$:

1. remove one block from the right, $\Delta c_i = 1$,
2. do not remove any block, $\Delta b_i = 1$,
3. remove one block on each side or the *spine*, $\Delta b_i = -1$,
4. remove one block from the left, $\Delta a_i = 1$.

Lexically ordered CSFs are obtained when the steps (1)–(4) are followed in the order given above. The top row, $i = n$, is uniquely defined by the total number of orbitals, n, total number of electrons, N and the target spin, S. If the ΔAC table is used to represent a configuration, the

corresponding Gel'fand table is simply represented by two binary strings of length n (number of orbitals).

By writing the row indices of the entries 1 (0) of the left (right) column of a ΔAC table into the second (first) column of a new table system, the Weyl table representation is obtained, (see Figure 6.5). The top row of the Gel'fand pattern, $[m]_n$, specifies the shape of the Weyl table, as it provides the number of boxes for each row of the Weyl table, i.e., m_{1n} boxes in the first row, m_{2n} boxes in the second row and so on. Due to the Pauli exclusion principle, $0 \leq m_{ij} \leq 2$, Weyl tables are restricted to two columns at most. This shape is then filled with *tokens* $1, 2, \ldots, n$, representing the occupied spatial orbitals. As shown in Figure 6.5 tokens are allowed to be *repeated* in a Weyl table. Tokens are non-decreasing from left to right in each row and increasing from top to bottom in each column. The practical interpretation of the Weyl table follows: it is an electronic configuration with a well defined spin (S), where the left column indicates the spin-up contributions (increasing the total spin by 1/2) and the right column the spin-down contribution (lower the total spin by 1/2). The value of each individual *token* represents the orbital in the given order.

There is a one-to-one correspondence between a Gelfand pattern, a Paldus table and a Weyl table specifying a CSF. An example of this correspondence is shown in Table 6.3 for a system of 2 electrons in 3 spatial orbitals for $S = 0$ and $S = 1$.

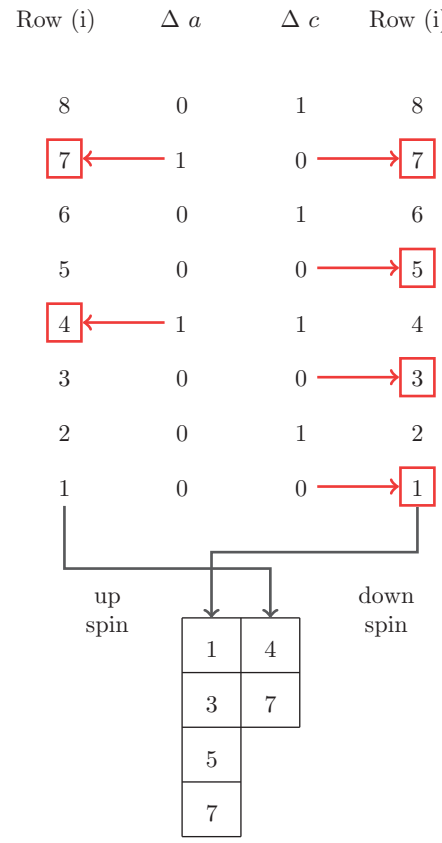
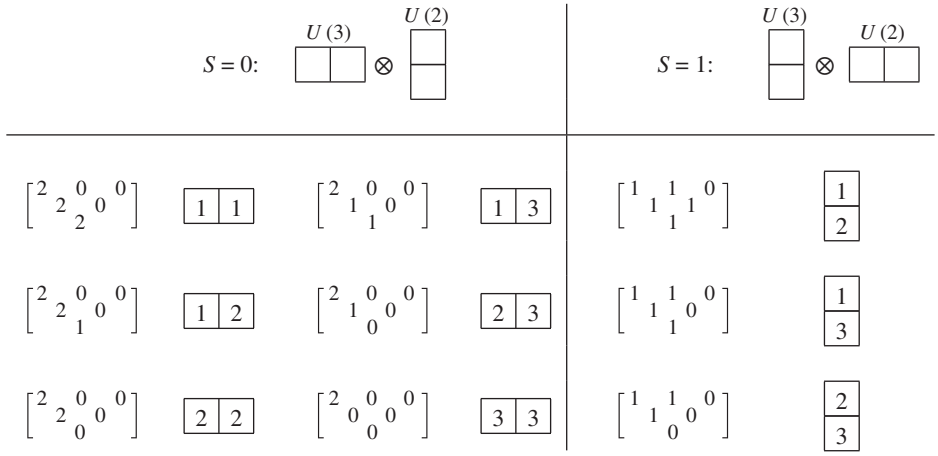


Figure 6.5 One-to-one mapping between ΔAC tables and Weyl table representing the same electronic configuration.

Table 6.3 Conjugate irreps of the spatial $U(3)$ and spin $U(2)$ group for the singlet, $S = 0$, and triplet, $S = 1$, states of 2 electrons in 3 spatial orbitals. The spatial U(n) irreps are determined by the spin-states states. All the CSFs belonging to this irrep are obtained by filling orbital "tokens" (1,2,3) with no repetitions in the same column in the Weyl table representation. The correspondence of Gel'fand pattern is also shown.



$\begin{bmatrix} 2 & 2 & 0 \\ & 2 & 0 \\ & & 0 \end{bmatrix}$	1 1	$\begin{bmatrix} 2 & 1 & 0 \\ & 1 & 0 \\ & & 0 \end{bmatrix}$	1 3	$\begin{bmatrix} 1 & 1 & 1 \\ & 1 & 1 \\ & & 0 \end{bmatrix}$	1 2
$\begin{bmatrix} 2 & 2 & 0 \\ & 1 & 0 \\ & & 0 \end{bmatrix}$	1 2	$\begin{bmatrix} 2 & 1 & 0 \\ & 0 & 0 \\ & & 0 \end{bmatrix}$	2 3	$\begin{bmatrix} 1 & 1 & 0 \\ & 1 & 0 \\ & & 0 \end{bmatrix}$	1 3
$\begin{bmatrix} 2 & 2 & 0 \\ & 0 & 0 \\ & & 0 \end{bmatrix}$	2 2	$\begin{bmatrix} 2 & 0 & 0 \\ & 0 & 0 \\ & & 0 \end{bmatrix}$	3 3	$\begin{bmatrix} 1 & 1 & 0 \\ & 1 & 0 \\ & & 0 \end{bmatrix}$	2 3

A practical example. Let us consider a system with three orbitals ($n = 3$) and four electrons ($N = 4$) coupled to a singlet spin ($S = 0$). For the top row (i = 3), we would have

$$a_3 = \frac{1}{2}N - S = 2 \tag{6.49}$$

$$b_3 = 2S = 0 \tag{6.50}$$

$$c_3 = n - a - b = 1 \tag{6.51}$$

Following the four rules given above all possible CSFs can be generated directly in lexical order as shown in Figure 6.6.

6.3.1.4 The Step-Vector

It has been discussed earlier that there are *four* possible actions to move from level i to the lower $(i-1)$ in the Paldus ΔAC table. These four different actions can be represented also by the more compact *step-vector*

$$d_i = 2\Delta a_i - \Delta c_i + 1. \tag{6.52}$$

The step vector can assume four integer values, 0, 1, 2 and 3, corresponding to one of the allowed actions. Table 6.4 summarizes the correspondence between the step-vector value and the variations Δa_i, Δb_i, Δc_i, ΔN_i and ΔS_i.

All basis functions of a chosen irrep of $U(n)$ can be generated by constructing all possible distinct step-vectors $|\mathbf{d}\rangle$ that lead to the same top-row of the Paldus table, Equations (6.42–6.44), with the restriction $S_i \geq 0, \forall\, i$. An example summarizing the connection between Paldus table, step-vector and Weyl table is given in Table 6.5.

6.3.2 The Graphical Unitary Group Approach (GUGA)

The graphical unitary group approach (GUGA) of Shavitt [13, 18] is based on the *step-vector* representation of CSFs and on the observation that there is a lot of repetition of possible rows in the Paldus arrays specifying the CSFs of a chosen irrep (see for instance the rows in Paldus tables of

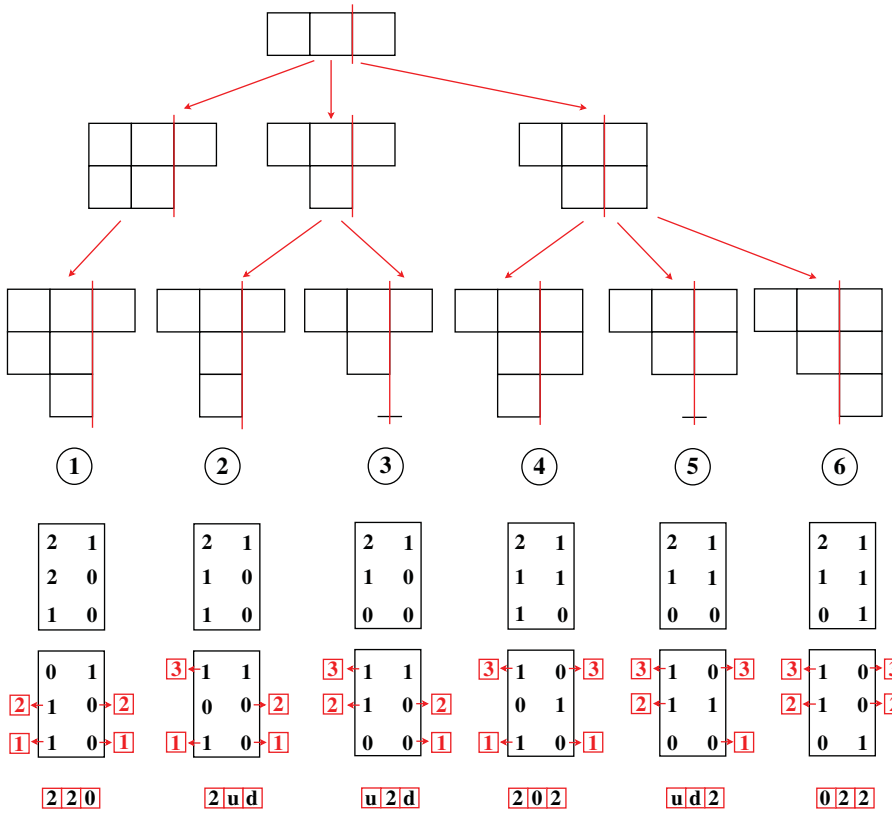


Figure 6.6 Example of the generation of a CSF basis in canonical lexical order for a system with 4 electrons in 3 orbitals coupled to a singlet state. The number of blocks on the left and right sides of each of the red "spines" represent the entries in the Paldus AC tables for each level (as in Figure 6.4). Paldus AC tables are translated into *variation-tables* and finally into Weyl's table (last row of tables).

Table 6.4 Possible values of the step-vector, d_i and corresponding variations in a_i, b_i, c_i, N_i and S_i.

d_i	Δa_i	Δb_i	Δc_i	ΔN_i	ΔS_i
0	0	0	1	0	0
1	0	1	0	1	1/2
2	1	−1	1	1	−1/2
3	1	0	0	2	0

Figure 6.6 above). Instead of storing all possible Paldus tables, Shavitt suggested to just list the possible sets of non-equivalent rows in a table, called the *distinct row table* (DRT). One example of DRT is given in Table 6.6. The number of all elements in this table is given by [13]

$$\begin{aligned} N_{DRT} &= (a+1)(c+1)\left(b+1+\frac{d}{2}\right) - \frac{d(d+1)(d+2)}{6} \\ &= \left(\frac{N}{2}-S+1\right)\left(n-\frac{N}{2}-S+1\right)\left(2S+1+\frac{d}{2}\right) - \frac{d(d+1)(d+2)}{6}, \end{aligned} \tag{6.53}$$

Table 6.5 Correspondence of Paldus table (a_i, b_i, N_i), step-vector (d_i) and Weyl table (right) representation of a CSF belonging to the $N = 8$, $n = 7$ and $S = 1$ irrep of $U(n)$. For each column in the Paldus table/step-vector, the current values of N_i and S_i are given.

orbital (i)	a_i	b_i	N_i	S_i	d_i
7	3	2	8	1	3
6	2	2	6	1	1
5	2	1	5	1/2	2
4	1	2	4	1	1
3	1	1	3	1/2	0
2	1	1	3	1/2	1
1	1	0	2	0	3
0	0	0	0	0	

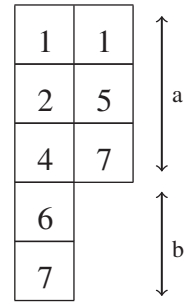


Table 6.6 Distinct row table for $n = 3$, $N = 4$ and $S = 0$.

a	b	c	i	j	d_0	d_1	d_2	d_3	u_0	u_1	u_2	u_3
2	0	1	3	1	2	0	3	4	-	-	-	-
2	0	0	2	2	0	0	0	5	1	0	0	0
1	1	0	2	3	0	5	0	6	0	0	1	0
1	0	1	2	4	5	0	6	7	0	0	0	1
1	0	0	1	5	0	0	0	8	4	3	0	2
0	1	0	1	6	0	8	0	0	0	0	4	3
0	0	1	1	7	8	0	0	0	0	0	0	4
0	0	0	0	8	-	-	-	-	7	6	0	5

with $d = \min(a, c) = \min(N/2 - S, n - N/2 - S)$, which is drastically smaller than the total number of corresponding CSFs (given by Eq. (6.8)) as seen in Figure 6.7.

Each row in DRT (see Table 6.6) is identified by a pair of indices (i, j) with $i = a_{ij} + b_{ij} + c_{ij}$ being the *level index*, with values from 0 to the number of space orbitals, n, and j being the *lexical row index* such that $j < j'$ if $a_{ij} > a'_{ij}$ or if $a_{ij} = a'_{ij}$ and $b_{ij} > b'_{ij}$.

A simple example. For a system with $n = 3$, $N = 4$ and $S = 0$ the DRT of Table 6.6 is derived.

Relations between distinct rows are indicated by the *downward*, d_{d_k}, and *upward*, u_{d_k}, *chaining indices*, with $d_k = 0, 1, 2, 3$ being the four possible step-values, mentioned in Section 6.3.1.4. These indices indicate the connection of a given lexical row to a neighboring level index after the action of a step-value, d. A zero entry indicates an invalid (non-existing) connection associated with the given step-value. From the DRT table any of the possible CSFs can be generated by connecting distinct rows linked by chaining indices.

This DRT table can be represented as a graph (Figure 6.8), where each distinct row is represented by a *vertex* (node) and non-zero chaining indices are indicated by *arcs* connecting the nodes. The arcs are labeled according to the *step number* (Eq. (6.52)). The vertices are labeled according to the lexical row index, j, starting at the unique *head* node at the top, which corresponds to the highest

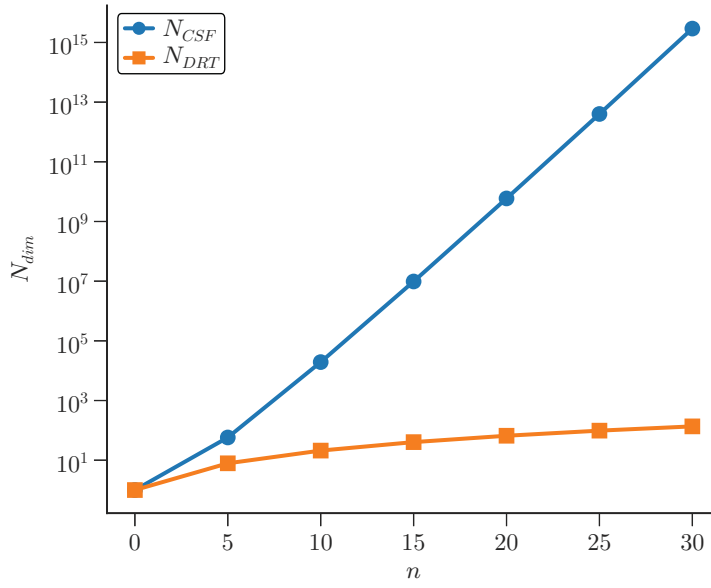


Figure 6.7 Number of total CSFs and entries of the distinct row table (DRT) for $S = 0$ and $N = n$ as a function of n.

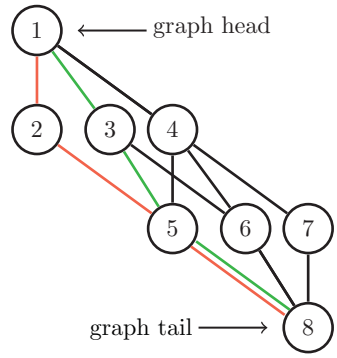


Figure 6.8 Graph representing the DRT of Table 6.6.

row (a, b, c), and ending at the second unique *tail* row, $(0, 0, 0)$. Vertices are arranged on a regular grid. Vertices with the same i-value of Table 6.6 are at the same level on this grid. The highest i-value is at the top and the lowest at the bottom. Vertices have also left–right order with respect to the a values, and vertices that share the same a value are further ordered (still horizontally) with respect to their b values.

CSFs are represented by *directed walks* through the graph. The green and orange lines in Figure 6.8 are two examples. Such a walk spans n arcs (number of orbitals) and visits one node at each hierarchical level. There is a direct correspondence between the Paldus table, Gel'fand patterns and directed walks on *Shavitt graphs* for representing all possible CSFs of a given electronic configuration.

A lengthy example. Consider a system with $n = 6$, $N = 5$ and $S = \frac{1}{2}$. Table 6.7 shows the 32 possible distinct rows for this system, while the total number of CSFs is $N_{CSF} = 210$. Notice that from the row with $j = 3$ we have generated only one *new* row that satisfies condition (4): remove

Table 6.7 Distinct row table for $n = 6$, $N = 5$ and $S = \frac{1}{2}$.

a	b	c	i	j	d_0	d_1	d_2	d_3	u_0	u_1	u_2	u_3
2	1	3	6	1	2	3	4	5	-	-	-	-
2	1	2	5	2	6	7	8	9	1	0	0	0
2	0	3	5	3	7	0	9	10	0	1	0	0
1	2	2	5	4	8	9	11	12	0	0	1	0
1	1	3	5	5	9	10	12	13	0	0	0	1
2	1	1	4	6	14	15	16	17	2	0	0	0
2	0	2	4	7	15	0	17	18	3	2	0	0
1	2	1	4	8	16	17	19	20	4	0	2	0
1	1	2	4	9	17	18	20	21	5	4	3	2
1	0	3	4	10	18	0	21	22	0	5	0	3
0	3	1	4	11	19	20	0	0	0	0	4	0
0	2	2	4	12	20	21	0	0	0	0	5	4
0	1	3	4	13	21	22	0	0	0	0	0	5
2	1	0	3	14	0	23	0	24	6	0	0	0
2	0	1	3	15	23	0	24	25	7	6	0	0
1	2	0	3	16	0	24	0	26	8	0	6	0
1	1	1	3	17	24	25	26	27	9	8	7	6
1	0	2	3	18	25	0	27	28	10	9	0	7
0	3	0	3	19	0	26	0	0	11	0	8	0
0	2	1	3	20	26	27	0	0	12	11	9	8
0	1	2	3	21	27	28	0	0	13	12	10	9
0	0	3	3	22	28	0	0	0	0	13	0	10
2	0	0	2	23	0	0	0	29	15	14	0	0
1	1	0	2	24	0	29	0	30	17	16	15	14
1	0	1	2	25	29	0	30	31	18	17	0	15
0	2	0	2	26	0	30	0	0	20	19	17	16
0	1	1	2	27	30	31	0	0	21	20	18	17
0	0	2	2	28	31	0	0	0	22	21	0	18
1	0	0	1	29	0	0	0	32	25	24	0	23
0	1	0	1	30	0	32	0	0	27	26	25	24
0	0	1	1	31	32	0	0	0	28	27	0	25
0	0	0	0	32	-	-	-	-	31	30	0	29

one block from the left. Conditions (1) is redundant (it would generate row (2,0,2) that already exists); condition (2) does not apply as it would raise i index; condition (3) is redundant as well. Obviously any row (a_{ij}, b_{ij}, c_{ij}) can only appear at level $i = a_{ij} + b_{ij} + c_{ij}$ and only between rows at levels $i \pm 1$ which are related to it by one of the four possible rules (1)–(4) given above. The graph corresponding to Table 6.7 is shown in Figure 6.9.

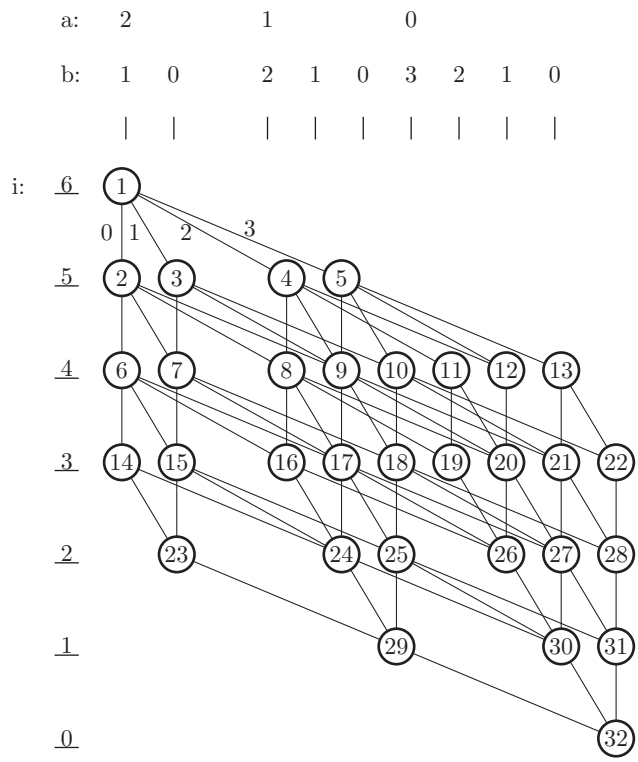


Figure 6.9 Graph representing the DRT of Table 6.7. The *a* and *b* values indicating the ordering in the graph are shown at the top and the level index *i* is indicated on the left. The different step-values *d* connecting the nodes are shown for the node 1 at the top.

With the ordering of the vertices according to their a and b values, the slope of each arc is in direct correspondence to the step-value d, connecting two vertices. Values of $d = 0$ correspond to vertical lines, and the slope of the other arcs increases with the step-value d.

A DRT table can be truncated by omitting distinct rows or by zeroing selected chaining indices. This approach is followed in generating restricted active space, RAS, or generalized active space, GAS, wave functions, as discussed in Section 6.6.

6.3.3 Evaluation of Non-Vanishing Hamiltonian Matrix Elements

Given the expression of the non-relativistic spin-free Hamiltonian operator, Eq. (1.18), a matrix element between two SDs or CSFs is given by

$$\langle m'|\hat{H}|m\rangle = \sum_{ij} h_{ij}\langle m'|\hat{E}_{ij}|m\rangle + \frac{1}{2}\sum_{ij,kl}(ij|kl)\langle m'|\hat{e}_{ij,kl}|m\rangle. \tag{6.54}$$

The matrix elements, $\langle m'|\hat{E}_{ij}|m\rangle$ and $\langle m'|\hat{e}_{ij,kl}|m\rangle$, are the *coupling coefficients* between two given SDs or CSFs, and h_{ij} and $(ij|kl)$ are the *electron repulsion integrals*. The coupling coefficients are independent of the orbital shape and depend only on the actual coupling between the SDs or CSFs, $|m'\rangle$ and $|m\rangle$, involved. Therefore, for a given set of MO integrals the problem of computing Hamiltonian matrix elements is simply reduced to the evaluation of these coupling coefficients. For SDs, the Slater–Condon rules apply to efficiently evaluate the coupling coefficients (see Section 6.2.1).

Unfortunately the Slater–Condon rules, cannot be applied for CSFs and the evaluation of the coupling coefficients between two CSFs is more involved, compared to SDs.

The graphical representation of CSFs is a powerful tool in evaluating these coupling coefficients, thanks to the formidable contribution of Paldus, Boyle and Shavitt [9, 19]. The strength of the graphical approach is the efficient identification and fast evaluation of non-vanishing coupling coefficients between two CSFs, $\langle m'|\hat{E}_{ij}|m\rangle$.

6.3.3.1 One-Body Coupling Coefficients

The excitation operators, $\hat{E}_{ij}$ are classified according to their indices. The diagonal $\hat{E}_{ii}$ elements are called *weight* (W) generators, elements $\hat{E}_{ij}$ with $i < j$ are referred to as *raising* (R) generators, and elements $\hat{E}_{ij}$ with $i > j$ are the *lowering* (L) generators. The excitation operators are called raising and lowering generators due to the fact that their effect on a given CSF is to increase or decrease their lexical order.

Opposed to Slater determinants, when a $\hat{E}_{ij}$ operator acts onto a CSF $|m\rangle$ yields a *linear combination* of CSFs, $|m'\rangle$

$$\hat{E}_{ij}|m\rangle = \sum_{m'} |m'\rangle\langle m'|\hat{E}_{ij}|m\rangle \tag{6.55}$$

with an electron moved from spatial orbital j to orbital i without changing the spin of the resulting states $|m'\rangle$.

A graphical representation of the Hamiltonian matrix elements is obtained by first identifying the two CSFs, $|m'\rangle$ and $|m\rangle$ in terms of their complete walks in the corresponding DRT graph (as in Figure 6.8). Next, a connection is made between the coupled CSFs and the coupling excitation operator of interest, for instance $\hat{E}_{ij}$. Shavitt [13] showed that for non-vanishing one-body coupling coefficients, the walks of the coupled CSFs, $|m\rangle$ and $|m'\rangle$, on the DRT graph must coincide outside the range (i,j), where i and j are the two MOs involved in the excitation via the $\hat{E}_{ij}$ operator. The range S_0, from $\min(i,j)$ to $\max(i,j)$, is referred to as the *range of the generator* $\hat{E}_{ij}$. The coupling coefficient will vanish if the walks in the DRT graph do not coincide outside the range (i,j). The two vertices in the DRT graph, related to orbital i and j represent the points of separation of the walks, and they are named *loop head* and *loop tail*. The two sections of overlapping walks are called *upper-* and *lower-walk*. The value of the coupling coefficient $\langle m'|\hat{E}_{ij}|m\rangle$ is independent of the the upper- and lower-walks. It only depends on the *shape of the loop* formed by the two graphs in the range (i,j) of the generator $\hat{E}_{ij}$. Figure 6.10 illustrates the graphical representation of the coupling coefficient together with the important elements of the graphs discussed above.

Considering the symmetry

$$\hat{E}_{ij}^{\dagger} = \hat{E}_{ji} \tag{6.56}$$

only the *raising* generators $\hat{E}_{ij}$ will be discussed in the following. Shavitt showed that for a general raising generator $\hat{E}_{ij}$, the condition of a non-vanishing coupling coefficient $\langle m'|\hat{E}_{ij}|m\rangle$ is that at each level

$$k = i, i+1, \ldots, j-1 \tag{6.57}$$

inside the loop, the vertices of the $|m'\rangle$ and $|m\rangle$ walks are related by the following relations

$$N'_k = N_k + 1 \quad \text{and} \quad S'_k = S_k \pm \frac{1}{2}, \quad \text{for} \quad k \in S_0, \tag{6.58}$$

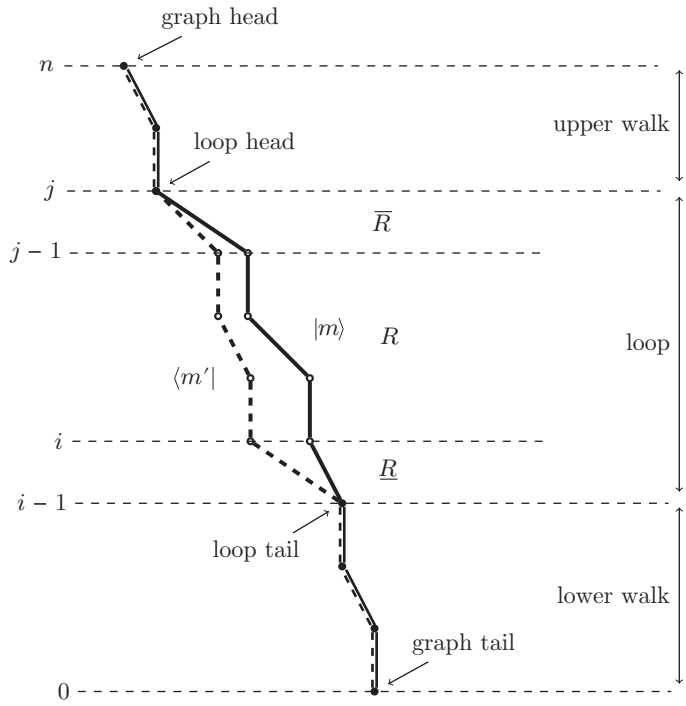


Figure 6.10 Graphical representation of a matrix element $\langle m'|\hat{E}_{ij}|m\rangle$ as a loop shape created by two CSFs $|m'\rangle$ and $|m\rangle$ on a Shavitt graph.

where N_k (N'_k) and S_k (S'_k) represent the cumulative occupation number and spin, respectively, up to the k level in the $|m\rangle$ ($|m'\rangle$) walk. The above conditions are equivalent to

$$2a'_k + b'_k = 2a_k + b_k + 1 \tag{6.59}$$

$$b'_k = b_k \pm 1. \tag{6.60}$$

in the Paldus representation of walks Eq.(6.45–6.46).

Based on the graphical approach, Shavitt [9] proved that matrix elements of the generators $\hat{E}_{ij}$ can be factorized in the product

$$\langle m'|\hat{E}_{ij}|m\rangle = \prod_{k=i}^{j} W(Q_k; d'_k, d_k, \Delta b_k, b_k), \tag{6.61}$$

where each term corresponds to a segment of the loop in the range S_0. The value b_k is the b value of $|m\rangle$ at level k. Additionally, $W(Q_k; d'_k, d_k, \Delta b_k, b_k)$ depends on the *segment shape* of the loop, determined by the *segment type* $Q_k = W$ (weight), R (raising) or L (lowering), the step values d'_k and d_k and $\Delta b_k = b_k - b'_k$. If the segment related to the $|m\rangle$ state is on the right, coincides or is on the left of the segment related to the $|m'\rangle$ state one will have raising (R), weight (W) or lowering (L) segments types, respectively. If the two segments of $|m'\rangle$ and $|m\rangle$ states close the loop at the top they are denoted $\bar{R}$ or $\bar{L}$. If they close the loop at the bottom they are denoted $\underline{R}$ or $\underline{L}$. The non-zero segment shapes for a raising (R) generator are shown in Figure 6.11.

In Table 6.8 non-vanishing one-electron coupling coefficients for the generator $\hat{E}_{ij}$ are given as functions of segment shape symbol, step-values and the b-value. They are expressed in terms of the

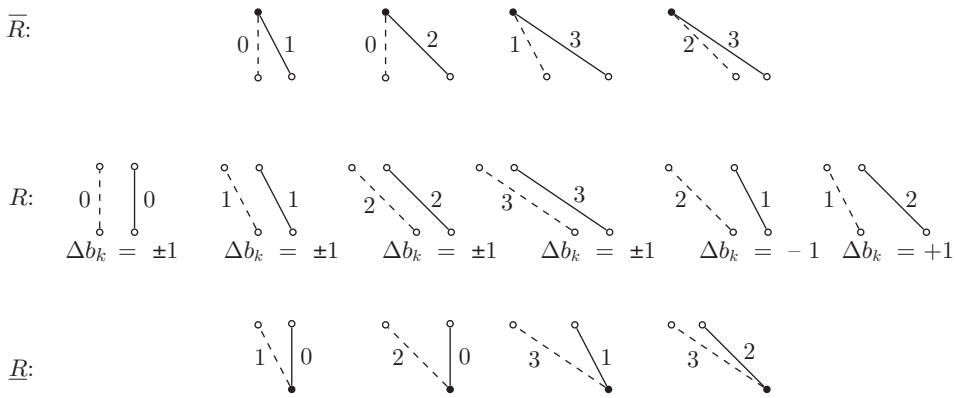


Figure 6.11 Non-zero segment shapes of a raising generator $\hat{E}_{ij}(i \geq j)$. The numbers next to the lines indicate the step-values d' and d. $\underline{R}(\bar{R})$ correspond to the loop tail (head) segments and R to shapes inside the generator range S_0. Δb_k indicates the possible difference of b'_k and b_k leading to non-zero matrix elements.

Table 6.8 Non-zero matrix contributions of the one-body operator $\hat{E}_{ij}$ in terms of the auxiliary functions Eq. (6.62, 6.63).

$d'd$	W
00	$-\mu$
11	$1-\mu$
22	$1-\mu$
33	$2-\mu$

$d'd$	$\bar{R}$	$\underline{L}$
01	1	1
02	1	1
13	$A(b,0,1)$	$A(b,2,1)$
23	$A(b,2,1)$	$A(b,0,1)$

$d'd$	$\underline{R}$	$\bar{L}$
10	1	1
20	1	1
31	$A(b,1,0)$	$A(b,0,1)$
32	$A(b,1,2)$	$A(b,2,1)$

	R		L	
$d'd$	$\Delta b=-1$	$\Delta b=+1$	$\Delta b=-1$	$\Delta b=+1$
00	1	1	1	1
11	-1	$C(b,0)$	$C(b,1)$	-1
12	$-1/(b+2)$	–	$1/(b+1)$	–
21	–	$1/b$	–	$-1/(b+1)$
22	$C(b,2)$	-1	-1	$C(b,1)$
33	-1	-1	-1	-1

following auxiliary functions

$$A(b,x,y) = \sqrt{\frac{b+x}{b+y}} \tag{6.62}$$

$$C(b,x) = \frac{\sqrt{(b+x-1)(b+x+1)}}{b+x}. \tag{6.63}$$

A simple example. Figure 6.12 shows the graphical representation of the matrix element $\langle 030300|\hat{E}_{25}|010320\rangle$, with $|m'\rangle$ and $|m\rangle$ in their step-vector representation. Each segment shape is obtained from Figure 6.11 and the corresponding value from Table 6.8 and the resulting matrix element as the product of them.

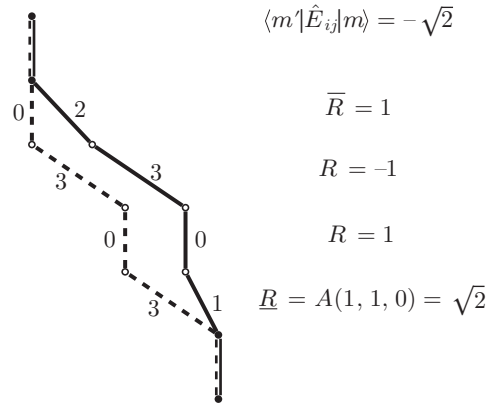


Figure 6.12 Graphical representation of the matrix element $\langle 030300|\hat{E}_{25}|010320\rangle = -\sqrt{2}$.

6.3.3.2 Two-Body Matrix Elements

The matrix elements of the two-body operators $\hat{e}_{ij,kl}$ are more involved than the one-body operators. Similar to the one-electron operators, the states $|m\rangle$ and $|m'\rangle$ must coincide outside the ranges (i,j) and (k,l) for $\langle m'|e_{ij,kl}|m\rangle$ to be non-zero. The form of the matrix element depends on the *overlap range* of the two ranges

$$S_1 = (i,j) \cap (k,l). \tag{6.64}$$

For non-overlapping ranges, $S_1 = \emptyset$, the matrix element just reduces to the product

$$\langle m'|e_{ij,kl}|m\rangle = \langle m'|E_{ij}E_{kl}|m\rangle = \langle m'|E_{ij}|m''\rangle\langle m''|E_{kl}|m\rangle, \tag{6.65}$$

where $|m''\rangle$ must coincide with $|m\rangle$ in the range (i,j) and with $|m'\rangle$ in range (k,l). The same rules and matrix elements as for one-body operators apply in this case.

For $S_1 \neq \emptyset$, we define the *non-overlap range*

$$S_2 = (i,j) \cup (k,l) - S_1, \tag{6.66}$$

where the same restrictions and matrix elements as for one-body operators apply. In the overlap range, S_1, different restrictions for the visited Paldus table vertices p apply for the matrix element to be non-zero. This depends on the type of the two generators involved and were worked out by Shavitt [18]. For two raising generators the following conditions apply

$$a'_p = a_p, \quad b'_p = b_p + 2, \quad c'_p = c_p - 2 \quad \text{for} \quad \Delta b_p = -2 \tag{6.67}$$

$$a'_p = a_p + 2, \quad b'_p = b_p + 2, \quad c'_p = c_p \quad \text{for} \quad \Delta b_p = +2 \tag{6.68}$$

$$a'_p = a_p + 1, \quad b'_p = b_p, \quad c'_p = c_p - 1 \quad \text{for} \quad \Delta b_p = 0; \tag{6.69}$$

for two lowering generators:

$$a'_p = a_p + 2, \quad b'_p = b_p + 2, \quad c'_p = c_p, \quad \text{for} \quad \Delta b_p = -2 \tag{6.70}$$

$$a'_p = a_p, \quad b'_p = b_p - 2, \quad c'_p = c_p + 2 \quad \text{for} \quad \Delta b_p = +2 \tag{6.71}$$

$$a'_p = a_p - 1, \quad b'_p = b_p, \quad c'_p = c_p + 2 \quad \text{for} \quad \Delta b_p = 0; \tag{6.72}$$

and for a mixed combination of raising and lowering generators:

$$a'_p = a_p - 1, \qquad b'_p = b_p + 2, \qquad c'_p = c_p - 1, \quad \text{for} \quad \Delta b_p = -2 \tag{6.73}$$

$$a'_p = a_p + 1, \qquad b'_p = b_p - 2, \qquad c'_p = c_p + 1 \quad \text{for} \quad \Delta b_p = +2 \tag{6.74}$$

$$a'_p = a_p, \qquad b'_p = b_p, \qquad c'_p = c_p \quad \text{for} \quad \Delta b_p = 0. \tag{6.75}$$

Drake and Schlesinger [20], Paldus and Boyle [19], Payne [21] and Shavitt and Paldus [18] were able to derive a scheme, where the two-body matrix elements can be computed as a product of *segment values* similar to the one-body case, Eq. (6.61)

$$\langle m'|\hat{e}_{ij,kl}|m\rangle = \prod_{p\in S_1} W(T_p, b_p) \sum_{x=0,1} \prod_{p\in S_2} W_x(T_p, b_p), \tag{6.76}$$

where S_1 and S_2 are the overlap and non-overlap ranges defined above Eq. (6.64, 6.66). $W(T_p, b_p)$ are the already defined single operator segment values (Table 6.8) and $W_x(T_p, b_p)$ are new segment values of the overlap range. There is a sum over two products in S_2, $x = 0$ corresponding to singlet coupled intermediate states, with a non-zero contribution if $\Delta b_p = 0, \forall p \in S_2$ and $x = 1$, corresponding to a triplet intermediate coupling. For further details interested readers are referred to Shavitt's paper [18].

6.4 Configuration Interaction Eigenvalue Problem

When replacing the continuous wave function with the discretized form

$$|\Psi\rangle = \sum_i C_i |i\rangle\,, \tag{6.77}$$

the Schrödinger equation reduces to a simple eigenvalue problem

$$\sum_i \langle j|\hat{H}|i\rangle\ C_i = E\ C_j\,, \tag{6.78}$$

and finding its solutions is equivalent to diagonalizing the $H_{ij} = \langle D_i|\hat{H}|D_j\rangle$ matrix, and obtaining its CI-eigenvectors, C^K, and eigenvalues, E^K. Slater determinants or CSFs may be used as the basis $\{D_i\}$ for building and diagonalizing the Hamiltonian matrix, with the CSFs preserving the spin symmetry as discussed earlier.

The Jacobi method [22], is an efficient *full diagonalization* technique, for finding all solutions of the eigenvalue problem. It is based on the assumption that the entire Hamiltonian matrix and eigenvalues for all states can be stored in memory. The computational cost for full diagonalizations increases as N^3_{conf} with N_{conf} being the dimension of the CI vector. As discussed in Section 6.1 the dimensionality of the CI space scales exponentially, and the Hamiltonian matrix is often too large to be stored in memory. In these cases the eigenvalue problem is not within reach for full diagonalization schemes, and other approaches are required. Also, in most cases a complete diagonalization is not needed, as only a few energetically lowest states are of interest. Efficient algorithms to obtain the lowest eigenvalues and the corresponding eigenvectors are crucial, considering the size of the CI vector.

Iterative methods have been recommended for state specific CI optimizations. The strength of these methods is that they do not require the entire list of Hamiltonian matrix elements and can rely on the contracted *σ-vector*

$$\boldsymbol{\sigma} = \mathbf{HC}\,. \tag{6.79}$$

This idea is the foundation of the Direct-CI approach, introduced by Roos in 1972 [6], and promoted by Handy and Knowles [4] and Olsen [7]. Methods to solve the eigenvalue problem for selected states are discussed in Section 6.4.1, while the Direct-CI algorithm is introduced in Section 6.4.2.

6.4.1 Iterative Methods

In this section we give a short overview on *iterative sparse eigensolvers*, here simply referred to as *iterative methods*, for finding specific solutions to the Schrödinger equation. Iterative methods are based on the principle of approaching a solution to Eq. (6.78) by repeated (iterative) application of $\hat{H}$, or a function of it, onto a trial vector $\mathbf{C}_0$. By projecting out any contribution from excited states to the state parametrized by $\mathbf{C}_0$, this procedure converges to the solution of Eq. (6.78).

The *power method* [23] is a simple iterative method, that solely relies on repeated application of $\hat{H}$. Consider a random initial vector $\mathbf{C}_0$ which can formally be decomposed into eigenvectors $\mathbf{v}_m$ of $\hat{H}$ as

$$\mathbf{C}_0 = \sum_m c_m \mathbf{v}_m \,, \tag{6.80}$$

with some coefficients c_m and eigenvalues E_m. Assume the eigenvectors are ordered such that E_0 has the highest absolute value. Then, applying $\hat{H}^n$ to $\mathbf{C}_0$ yields

$$\begin{aligned} \hat{H}^n \mathbf{C}_0 &= \sum_m E_m^n c_m \mathbf{v}_m \\ &= E_0^n \left(c_0 \mathbf{v}_0 + \sum_{m>0} c_m \left(\frac{E_m}{E_0} \right)^n \mathbf{v}_m \right) \to E_0^n c_0 \mathbf{v}_0 (n \to \infty, \text{normalized}) \,. \end{aligned} \tag{6.81}$$

By shifting the Hamiltonian such that the ground state energy has the highest absolute value, this scheme, when normalizing the vector, converges to the ground state CI vector. The convergence of this scheme can be very slow depending on the fractions $\frac{E_m}{E_0}$, and even though $\hat{H}^n \mathbf{C}_0$ is computed for a wide range of n, only the last one is utilized, discarding all intermediate results. For these reasons, while the simple scheme of repeatedly applying the transformation $\mathbf{C} \mapsto \hat{H}\mathbf{C}$ is a valid approach, it has little practical relevance for solving Eq. (6.78).

6.4.1.1 Lanczos Algorithm

In the *Lanczos algorithm* [24] the intermediate results

$$\mathbf{C}_n = \hat{H}^n \mathbf{C}_0 \tag{6.82}$$

are utilized by implementing the ground state search as a Krylov-subspace procedure [25]. In the Lanczos procedure, at each iteration, the following ansatz space is considered

$$K = \text{span}\{\mathbf{C}_0, \mathbf{C}_1, \mathbf{C}_2 \ldots\} \tag{6.83}$$

spanned by the initial guess $\mathbf{C}_0$ and all previous intermediate results, $\mathbf{C}_n$, obtained by $\mathbf{C} \mapsto H\mathbf{C}$. The vectors $\mathbf{C}_n$ are not orthogonal by default; hence, the basis of K is orthonormalized in each iteration, leading to a new basis $\{\mathbf{b}_i\}$ of K

$$K = \text{span}\{\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2 \ldots\} \,. \tag{6.84}$$

The application of $\hat{H}$ to $\mathbf{b}_n$ expands the ansatz space K by a new basis vector, $\mathbf{C} = H\mathbf{b}_n$. By this scheme an orthonormal basis of K, $\{\mathbf{b}_i\}$, is iteratively obtained, with each new vector

$$\mathbf{b}_{n+1} \in \text{span}\{\mathbf{b}_0,\ \mathbf{b}_1\ \ldots\ \mathbf{b}_n,\ H\mathbf{b}_n\} \tag{6.85}$$

$$\hat{H} = \begin{bmatrix} H_{11} & H_{12} & 0 & \cdots & 0 \\ H_{21} & \ddots & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & H_{(n-1)n} \\ 0 & \cdots & & H_{n(n-1)} & H_{nn} \end{bmatrix}$$

Figure 6.13 Hamiltonian matrix in tridiagonal form.

orthogonalized against the existing K-space basis vectors. In the case that $\hat{H}\mathbf{b}_n \in K$, $\mathbf{b}_{n+1}$ is chosen as a random vector orthogonal to K instead, as under those circumstances orthogonalization is not possible. The full scheme is summarized in Algorithm 1.

Given the hermiticity of $\hat{H}$ (symmetric for real-valued matrix elements), a particular property of the Lanczos method is that $\hat{H}$ is *tridiagonal* on K (see Figure 6.13). This follows directly from the fact that

$$H_{ij} = \mathbf{b}_i^{\mathbf{T}} \hat{H} \mathbf{b}_j = 0 \qquad \text{if} \qquad j < i - 1\,, \tag{6.86}$$

for $\mathbf{b}_i$ is orthogonalized against all previous vectors $\mathbf{b}_j$, including $H\mathbf{b}_j$ for all $j < i - 1$. From the hermiticity of $\hat{H}$ it also follows that

$$H_{ij} = \mathbf{b}_j^{\mathbf{T}} \hat{H} \mathbf{b}_i = 0 \qquad \text{if} \qquad j < i - 1\,, \tag{6.87}$$

which implies that the matrix elements H_{ij} of $\hat{H}$ on K are non-zero only for $|i - j| \leq 1$

$$H_{ij} \neq 0 \Rightarrow \ |i - j| \leq 1\,. \tag{6.88}$$

This particular property makes the Lanczos algorithm inexpensive in terms of memory requirements, as only the last two vectors $\mathbf{b}_{n-1}$ and $\mathbf{b}_n$ are required to obtain the new orthogonal vector $\mathbf{b}_{n+1}$.

6.4.1.2 Davidson Algorithm

The Davidson algorithm [1, 2], as for the Lanczos approach, is an iterative method, based on the repeated application of $\hat{H}$ onto a trial vector. The main difference with the Lanczos method lies in how the basis vector is updated.

Algorithm 1 Lanczos Algorithm

- Choose an initial state $\mathbf{b}_0$ and a maximum iteration number m

for $n = 0, \ldots, m$ **do**

- Construct the space $K = \text{span}\{\mathbf{b}_0, \ldots \mathbf{b}_n\}$
- Obtain $\mathbf{C} = \hat{H}\mathbf{b}_n$

if $\mathbf{C} \in K$ **then**

- Set b_{n+1} to a random vector orthogonal to K

else

- Orthogonalize $\mathbf{C}$ against K, yielding $\mathbf{b}_{n+1}$

end if

end for

The Davidson method can be derived from the *quasi-Newton method*, applied to solve Eq. (6.78) [26]. We decompose the Hamiltonian operator into a dominant part and a correction

$$\hat{H} = \hat{H}_0 + \hat{H}_1, \tag{6.89}$$

The updated CI vector, $\mathbf{C'}$ is also decomposed into an initial guess, $\mathbf{C}$, and a correction $\mathbf{b}$

$$\mathbf{C}' = \mathbf{C} + \mathbf{b}\,, \tag{6.90}$$

The correction $\mathbf{b}$ can be separated into a component $\mathbf{b}_{\parallel} = \lambda\mathbf{C}$ parallel to $\mathbf{C}$ and a component $\mathbf{b}_{\perp}$ perpendicular to $\mathbf{C}$

$$\mathbf{b} = \lambda\mathbf{C} + \mathbf{b}_{\perp}\,. \tag{6.91}$$

As rescaling $\mathbf{C}'$ via the parallel component of $\mathbf{b}$ does not influence its property of being or not being a solution of Eq. (6.78), the contribution from $\mathbf{b}_{\parallel}$ can always be eliminated and the update simplified as

$$\mathbf{C}' = \mathbf{C} + \mathbf{b}_{\perp}\,, \tag{6.92}$$

The eigenvalue problem is thus expressed as

$$(\hat{H}_0 + \hat{H}_1)(\mathbf{C} + \mathbf{b}_{\perp}) = E(\mathbf{C'})(\mathbf{C} + \mathbf{b}_{\perp})\,. \tag{6.93}$$

If the corrections are small and second order terms in H_1 and $\mathbf{b}$ are neglected, the eigenvalue problem becomes

$$\mathbf{b}_{\perp} = -(\hat{H}_0 - E(\mathbf{C}))^{-1}(H - E(\mathbf{C'}))\mathbf{C}\,, \tag{6.94}$$

where we have defined

$$E(\mathbf{C}) = \frac{\mathbf{C}H\mathbf{C}}{\mathbf{C}\mathbf{C}}\,. \tag{6.95}$$

The right hand side still depends on $\mathbf{b}$ via $E(\mathbf{C}')$. However, by considering that $\mathbf{C}$ and $\mathbf{b}_{\perp}$ are orthogonal one can solve for $E(\mathbf{C'})$

$$0 = \mathbf{C}\mathbf{b}_{\perp} = -\mathbf{C}(H_0 - E(\mathbf{C}))^{-1}(\hat{H} - E(\mathbf{C'}))\mathbf{C}\,. \tag{6.96}$$

leading to

$$E(\mathbf{C'}) = \frac{\mathbf{C}(H_0 - E(\mathbf{C}))^{-1}H\mathbf{C}}{\mathbf{C}(H_0 - E(\mathbf{C}))^{-1}\mathbf{C}} \tag{6.97}$$

Inserting Eq. (6.97) into Eq. (6.94) the explicit expression for the *quasi-Newton* step is obtained

$$\mathbf{b} = -(H_0 - E(\mathbf{C}))^{-1}\left((\hat{H} - E(\mathbf{C})) - \frac{\mathbf{C}\cdot(H_0 - E(\mathbf{C}))^{-1}(H - E(\mathbf{C}))\mathbf{C}}{\mathbf{C}\cdot(H_0 - E(\mathbf{C}))^{-1}\mathbf{C}}\right)\mathbf{C}\,. \tag{6.98}$$

This approach relies on the assumption that higher order terms, H_1 and $\mathbf{b}_{\perp}$ are negligible, and therefore that the initial vector $\mathbf{C}$ is close to the actual solution $\mathbf{C}_{\text{eigen}}$. Like with the power method, one relies on an iterative scheme, that requires repeated application of Eq. (6.98) to gradually converge to $\mathbf{C}_{\text{eigen}}$, as each iteration decreases the required correction $\mathbf{b}_{\perp}$ and therefore improves the quality of the expansion, yielding a better result in the next iteration. The vector $\mathbf{C}'$ is not necessarily normalized after each iteration.

The strength of this method rests on the good partition (Eq. (6.89)) [27], as both, the effect of $\hat{H}_1$ cannot be too large, which would hinder convergence, and the inverse of $(H_0 - E(\mathbf{C}))$ must be easily computable. Under these circumstances, the quasi-Newton method is an efficient alternative

to the simple power method, at comparable cost per iteration, while potentially converging much faster [26].

At convergence, it is $E(\mathbf{C}) = E(\mathbf{C}')$ and one additional approximation may be employed in Eq. (6.94), namely replacing $E(\mathbf{C}')$ with $E(\mathbf{C})$, leading to the *Davidson step*

$$\mathbf{b} = -(H_0 - E(\mathbf{C}))^{-1}(\hat{H} - E(\mathbf{C}))\mathbf{C} \,. \tag{6.99}$$

The *Davidson method* is one of the most frequently applied methods for treating the CI problem in quantum chemistry. While at comparable cost per iteration, in contrast to the quasi-Newton method, the Davidson algorithm is less sensitive to H_0 being close to H, but also does not necessarily improve when $\hat{H}_0 \to \hat{H}$. A common choice for H_0 is the diagonal part of H [1], which minimizes the effort required to invert $(\hat{H}_0 - E(\mathbf{C}))$.

As the residual vector

$$\mathbf{r} = (\hat{H} - E(\mathbf{C}))\mathbf{C} \tag{6.100}$$

is computed to perform each step, an estimate of convergence is readily obtained from the norm of $\mathbf{r}$.

Typically, the Davidson method is carried out building an ansatz space from the results of previous iterations, similar to the Lanczos method. As the correction vectors are not orthogonal by default, an additional orthogonalization step has to be performed after each iteration. As each iteration increases the dimension of K by one, it can be useful to truncate the ansatz space to some maximum dimension to keep the orthogonalization step feasible. The Davidson algorithm is summarized in Algorithm 2.

Algorithm 2 Davidson Algorithm

- Choose an initial state $\mathbf{b}_0$ and a maximum iteration number m
for $n = 0, \dots, m$ **do**
  - Construct the space $K = \text{span}\{\mathbf{b}_0, \dots, \mathbf{b}_n\}$
  - Obtain the ground state $\mathbf{C}$ of $\hat{H}$ in the K-space
  if $|(H - E(\mathbf{C}))\mathbf{C}|$ is below tolerance **then**
    - Exit
  end if
  - Obtain the next vector $\mathbf{b}_{n+1}$ as
    $\mathbf{b}_{n+1} = -\left(H_0 - E(\mathbf{C})\right)^{-1}(H - E(\mathbf{C}))\,\mathbf{C}$
end for

6.4.2 Direct-CI Algorithm

Let us write the CI expansion as

$$|0\rangle = \sum_{I_\alpha I_\beta} C(I_\alpha, I_\beta)|\alpha(I_\alpha)\beta(I_\beta)\rangle. \tag{6.101}$$

This form of the CI expansion highlights the factorization of Slater determinants into α (I_α) and β (I_β) strings and the possibility to recast the CI vector into a rectangular matrix, $C(I_\alpha, I_\beta)$. Using this form of the CI vector the σ vector can also be expressed as a rectangular matrix

$$\sigma(I_\alpha, I_\beta) = \sum_{J_\alpha J_\beta} \langle\alpha(I_\alpha)\beta(I_\beta)|\hat{H}|\alpha(J_\alpha)\beta(J_\beta)\rangle C(J_\alpha, J_\beta). \tag{6.102}$$

Slater–Condon rules are used to evaluate the Hamiltonian matrix elements. Non-zero terms are multiplied by the corresponding $C(I_\alpha, I_\beta)$ and the contribution added to the sigma vector. While this simple scheme is a valid approach, it is highly inefficient as it scales as N^2_{conf}. The direct-CI approach is based on the realization that Slater determinants are coupled via single and double excitations, and for a given Slater determinant it is possible to know a priori the entire list of coupled determinants, while avoiding the explicit evaluation of the vanishing Hamiltonian matrix elements. Although still quite costly, this method scales as the number of non-vanishing matrix elements, significantly reduced with respect to the explicit evaluation of all Hamiltonian matrix elements. The *minimal operation-count (MOC)* variant of the Direct-CI was introduced by Olsen *et al.* in 1988 [7] and it is entirely based on the possibility to separate the α and β contributions in the evaluation of the sigma vector. For simplicity we rewrite the Hamiltonian operator as

$$\hat{H} = \sum_{pq} k_{pq}\hat{E}_{pq} + \frac{1}{2}\sum_{pq,rs}(pq|rs)\hat{E}_{pq}\hat{E}_{rs}, \tag{6.103}$$

where

$$k_{pq} = h_{pq} - \frac{1}{2}\sum_{r}(pr|rq). \tag{6.104}$$

The one- and two-electron terms of the σ-vector can be separated as

$$\sigma(I_\alpha, I_\beta) = \sigma(I_\alpha, I_\beta)^{(1)} + \sigma(I_\alpha, I_\beta)^{(2)} \tag{6.105}$$

with

$$\sigma(I_\alpha, I_\beta)^{(1)} = \sum_{pq}\sum_{J_\alpha J_\beta} k_{pq}\langle\alpha(I_\alpha)\beta(I_\beta)|\hat{E}_{pq}|\alpha(J_\alpha)\beta(J_\beta)\rangle C(J_\alpha, J_\beta), \tag{6.106}$$

$$\sigma(I_\alpha, I_\beta)^{(2)} = \frac{1}{2}\sum_{pqrs}\sum_{J_\alpha J_\beta}(pq|rs)\langle\alpha(I_\alpha)\beta(I_\beta)|\hat{E}_{pq}\hat{E}_{rs}|\alpha(J_\alpha)\beta(J_\beta)\rangle C(J_\alpha, J_\beta). \tag{6.107}$$

Considering the separation of the α and β component of the excitation operator (Chapter 1, Eq. (1.19)), the $\sigma^{(1)}$ term can be further decomposed into the α and the β contribution

$$\sigma(I_\alpha, I_\beta)^{(1)} = \sigma(I_\alpha, I_\beta)^{(1\alpha)} + \sigma(I_\alpha, I_\beta)^{(1\beta)} \tag{6.108}$$

with

$$\sigma(I_\alpha, I_\beta)^{(1\alpha)} = \sum_{pq}\sum_{J_\alpha J_\beta} k_{pq}\langle\alpha(I_\alpha)|\hat{E}^\alpha_{pq}|\alpha(J_\alpha)\rangle\delta_{\beta(I_\beta)\beta(J_\beta)}C(J_\alpha, J_\beta) \tag{6.109}$$

$$= \sum_{pq}\sum_{J_\alpha} k_{pq}\langle\alpha(I_\alpha)|\hat{E}^\alpha_{pq}|\alpha(J_\alpha)\rangle C(J_\alpha, I_\beta). \tag{6.110}$$

Introducing the matrix representation of the one-electron operator

$$\hat{k}^\alpha_{I_\alpha J_\alpha} = \sum_{pq} k_{pq}\langle\alpha(I_\alpha)|\hat{E}^\alpha_{pq}|\alpha(J_\alpha)\rangle \tag{6.111}$$

we finally obtain

$$\sigma(I_\alpha, I_\beta)^{(1\alpha)} = \sum_{J_\alpha}\hat{k}^\alpha_{I_\alpha J_\alpha} C(J_\alpha, I_\beta). \tag{6.112}$$

The matrix in Eq. (6.111) is rather sparse. For a given I_α there are only $[N_\alpha(n - N_\alpha) + 1]$ non-vanishing $\hat{k}^\alpha_{I_\alpha J_\alpha}$ terms, as that is the number of coupling J_α terms via the single excitation operator (n is the total number of orbitals, N_α the number of α electrons). Additionally the computed $\hat{k}^\alpha_{I_\alpha J_\alpha}$ terms are the same for any choice of I_β string. An identical argument is carried

out for the $\sigma(I_\alpha, I_\beta)^{(1\beta)}$. The evaluation of the $\sigma(I_\alpha, I_\beta)^{(1)}$ is therefore reduced to two matrix-matrix multiplications, $\mathbf{k} \times \mathbf{C}$, with a total operation count (to the leading term) of

$$[N_\alpha(n - N_\alpha) + N_\beta(n - N_\beta)]N_{det} \tag{6.113}$$

The two-electron contribution can be divided into three terms: one term involving two α excitations, one involving two β excitations, and a third one involving mixed excitations

$$\hat{H}^{(2)} = \hat{H}^{(2)}_{\alpha\alpha} + \hat{H}^{(2)}_{\beta\beta} + \hat{H}^{(2)}_{\alpha\beta} \tag{6.114}$$

$$H^{(2)}_{\alpha\alpha} = \frac{1}{2}\sum_{pqrs}(pq|rs)\hat{E}^{\alpha}_{pq}\hat{E}^{\alpha}_{rs} \tag{6.115}$$

$$H^{(2)}_{\beta\beta} = \frac{1}{2}\sum_{pqrs}(pq|rs)\hat{E}^{\beta}_{pq}\hat{E}^{\beta}_{rs} \tag{6.116}$$

$$H^{(2)}_{\alpha\beta} = \sum_{pqrs}(pq|rs)\hat{E}^{\alpha}_{pq}\hat{E}^{\beta}_{rs} \tag{6.117}$$

The contribution to the σ-vector thus is

$$\sigma(I_\alpha, I_\beta)^{(2)}_{\alpha\alpha} = \frac{1}{2}\sum_{pqrs}\sum_{J_\alpha}(pq|rs)\langle\alpha(I_\alpha)|\hat{E}^{\alpha}_{pq}\hat{E}^{\alpha}_{rs}|\alpha(J_\alpha)\rangle C(J_\alpha, I_\beta) \tag{6.118}$$

$$\sigma(I_\alpha, I_\beta)^{(2)}_{\beta\beta} = \frac{1}{2}\sum_{pqrs}\sum_{J_\beta}(pq|rs)\langle\beta(I_\beta)|\hat{E}^{\beta}_{pq}\hat{E}^{\beta}_{rs}|\beta(J_\beta)\rangle C(I_\alpha, J_\beta) \tag{6.119}$$

$$\begin{aligned}\sigma(I_\alpha, I_\beta)^{(2)}_{\alpha\beta} = &\sum_{pqrs}\sum_{J_\alpha J_\beta}(pq|rs)\langle\alpha(I_\alpha)|\hat{E}^{\alpha}_{pq}|\alpha(J_\alpha)\rangle \\ &\times\langle\beta(I_\beta)|\hat{E}^{\beta}_{rs}|\beta(J_\beta)\rangle C(J_\alpha, J_\beta)\end{aligned} \tag{6.120}$$

As for the one-electron contribution, a matrix representation of the two-electron operators may be introduced

$$G^{\alpha\alpha}_{I_\alpha J_\alpha} = \frac{1}{2}\sum_{pqrs}(pq|rs)\langle\alpha(I_\alpha)|\hat{E}^{\alpha}_{pq}\hat{E}^{\alpha}_{rs}|\alpha(J_\alpha)\rangle \tag{6.121}$$

and the σ-vector contribution rewritten as

$$\sigma(I_\alpha, I_\beta)^{(2)}_{\alpha\alpha} = \sum_{J_\alpha} G^{\alpha\alpha}_{I_\alpha J_\alpha} C(J_\alpha, I_\beta) \tag{6.122}$$

$$\sigma(I_\alpha, I_\beta)^{(2)}_{\beta\beta} = \sum_{J_\beta} G^{\beta\beta}_{I_\beta J_\beta} C(I_\alpha, J_\beta) \tag{6.123}$$

For a given $\alpha(I_\alpha)$ string the number of non-vanishing $G^{\alpha\alpha}_{I_\alpha J_\alpha}$ is $\frac{1}{4}[N^2_\alpha(n - N_\alpha)^2]$ (to the leading terms). Therefore the operation count for the $\sigma(I_\alpha, I_\beta)^{(2)}_{\alpha\alpha}$ and $\sigma(I_\alpha, I_\beta)^{(2)}_{\beta\beta}$ is

$$\frac{1}{4}[N^2_\alpha(n - N_\alpha)^2 + N^2_\beta(n - N_\beta)^2]N_{det} \tag{6.124}$$

For the mixed term $\sigma(I_\alpha, I_\beta)^{(2)}_{\alpha\beta}$ the following intermediate matrices are introduced

$$G^{\beta}_{I_\beta J_\beta}[pq] = \sum_{rs}(pq|rs)\langle\beta(I_\beta)|\hat{E}^{\beta}_{rs}|\beta(J_\beta)\rangle \tag{6.125}$$

$$D^{\alpha}_{I_\alpha J_\beta}[pq] = \sum_{J_\alpha}\langle\alpha(I_\alpha)|\hat{E}^{\alpha}_{pq}|\alpha(J_\alpha)\rangle C(J_\alpha J_\beta) \tag{6.126}$$

and the $\sigma(I_\alpha, I_\beta)^{(2)}_{\alpha\beta}$ term can be rewritten as

$$\sigma(I_\alpha, I_\beta)^{(2)}_{\alpha\beta} = \sum_{pq} \sum_{J_\beta} G^{\beta}_{I_\beta J_\beta}[pq] D^{\alpha}_{I_\alpha J_\beta}[pq] \tag{6.127}$$

For a given $\alpha(I_\alpha)$ there are only $[N_\alpha(n - N_\alpha) + 1]$ non-vanishing $D^{\alpha}_{I_\alpha J_\beta}[pq]$ with variable pq pairs. This operation count is independent of the $\beta(J_\beta)$ value and is to be multiplied for the total number of determinants, N_{det}. The number of non-vanishing $G^{\beta}_{I_\beta J_\beta}[pq]$ depends on the β-strings, namely for each $\beta(I_\beta)$ there are $[N_\beta(n - N_\beta) + 1]$ non-vanishing G terms. Therefore the total operation count for $\sigma(I_\alpha, I_\beta)^{(2)}_{\alpha\beta}$ is to the leading term

$$N_\alpha N_\beta (n - N_\alpha)(n - N_\beta) N_{det} \tag{6.128}$$

Thus the total operation count for the three two-electron components of the σ-vector is

$$\left[\frac{1}{4} N_\alpha^2 (n - N_\alpha)^2 + \frac{1}{4} N_\beta^2 (n - N_\beta)^2 + N_\alpha N_\beta (n - N_\alpha)(n - N_\beta)\right] N_{det} \tag{6.129}$$

Graphical methods to label and order Slater determinants as discussed in Section 6.2 further simplify the mapping between coupled α- and β-strings, increasing the overall efficiency of the Direct-CI algorithm.

6.5 The CASSCF Method

The Hartree–Fock (HF) method has contributed to the popularity of quantum chemistry in electronic structure predictions, due to its ability to provide accurate zeroth order reference wave functions for closed-shell molecular systems in their ground state geometries. However, the HF method neglects electron correlation (except for the spin) and fails in situations where this is important, such as in bond formation or breaking processes. In these situations single configurational wave functions, such as the HF wave function, are not sufficient and, more than one configuration is necessary to obtain a qualitatively correct wave function. Multi-configurational (MC) approaches have been developed as an extension to the Hartree–Fock method to circumvent its limitations for MC molecular systems. MC self-consistent field (MCSCF) approaches have played a central role in the field of MC methods. In MCSCF both CI coefficients and molecular orbitals are variationally optimized. The optimization of the molecular orbitals lifts the bias introduced by the initial conditions. The purpose of the MCSCF method, also considering its *state-averaged* formulation, is to obtain qualitatively correct electronic wave functions for all states considered along a given potential energy surface. Correlation effects not described by the MCSCF wave function are generally recovered by subsequent MRCI or PT2 treatments that use the MCSCF wave function as zeroth order approximation.

A Two-Configuration Wave Function for the Dissociation of H_2

The failure of the restricted Hartree–Fock (RHF) method in describing the dissociation of the H_2 molecule is used here as motivation for more advanced MC approaches. We construct symmetry adapted orthonormal molecular orbitals from a minimal basis set

$$\sigma_g = N_g(1s_A + 1s_B) \tag{6.130}$$

$$\sigma_u = N_u(1s_A - 1s_B), \tag{6.131}$$

where σ_g and σ_u are the bonding and anti-bonding MOs, respectively. N_g and N_u are the normalization factors and $1s_A$ and $1s_B$ the atomic orbitals centered at atom A and B of the hydrogen molecule. The ground state RHF wave function is represented by the following single Slater determinant

$$\Psi = (\sigma_g)^2 = \sqrt{\frac{1}{2}}|\sigma_g(1)\alpha(1)\sigma_g(2)\beta(2)| \tag{6.132}$$

Expanding the Slater determinant one obtains

$$\Psi = N_g^2[1s_A(1)1s_A(2) + 1s_A(1)1s_B(2) + 1s_B(1)1s_A(2) + 1s_B(1)1s_B(2)]\theta(1,2) \tag{6.133}$$

Terms $1s_A(1)1s_B(2)$ and $1s_B(1)1s_A(2)$ describe the homolytic dissociation of the molecule, with one electron in $1s_A$ and one in $1s_B$. Terms $1s_A(1)1s_A(2)$ and $1s_B(1)1s_B(2)$ represent the unphysical and undesired ionic dissociation to H_A^- and H_B^+. The function $\theta(1,2)$ is a function of the spin. The energy expression at dissociation, associated to the single-determinantal wave function is

$$E^{\infty} = 2E(H) + \frac{1}{2}(1s_A1s_A|1s_A1s_A) \tag{6.134}$$

The term $2E(H)$ is twice the value of the atomic energy and, it is the only value expected to remain in the energy expression for the two non-interacting hydrogen atoms. The above energy expression, however, contains the extra term $(1s_A1s_A|1s_A1s_A)$. This term derives from the ionic configurations and it is positive, unphysically raising the total energy for the system at dissociation.

To overcome this problem the following two-configuration wave function is introduced

$$|0\rangle = C_1\Psi_1 + C_2\Psi_2, \tag{6.135}$$

with Ψ_1 defined as in Eq. (6.132) and Ψ_2 represented by

$$\Psi_2 = (\sigma_u)^2 = \sqrt{\frac{1}{2}}|\sigma_u(1)\alpha(1)\sigma_u(2)\beta(2)| \tag{6.136}$$

At dissociation $C_1 = -C_2$ and the ionic contributions cancel out, revealing the correct wave function and energy. The two-configuration wave function for the dissociation of the H_2 molecule is a simple example of a MC wave function.

In the early days of MCSCF, small CI expansions were constructed by a manual selection of the relevant determinants and, CI coefficients and molecular orbitals were simultaneously optimized. The manual selection of electronic configurations, however, was rather problematic as it required experience based on a trial-and-error approach and could lead to biased results, as missing configurations could be important for the chemical system investigated. The *complete active space self-consistent field* (CASSCF) approach [28–31] is a special form of MCSCF in which the configuration interaction expansion is chosen to be *complete* inside a user-specific orbital subset, the *active space*. In CASSCF, orbitals are divided into five groups. *Frozen* orbitals are doubly occupied in all configurations of the CI expansion and their shape is fixed to the one obtained in the preceding HF optimization, assuming that HF is the method used to generate the starting orbitals. *Inactive* orbitals are doubly occupied in all electronic configurations of the CI expansion. Contrarily to the frozen orbitals, the inactive orbital coefficients are optimized during the MCSCF procedure. *Active* orbitals represent the basis in which a *full-CI* expansion is generated. Thus, their occupation can vary in the range $[0-2]$ in the electronic configurations of the CI expansion. *Virtual* (or *secondary*) orbitals are the empty analogs of the inactive orbitals. These orbitals are empty in all electronic configurations of the CI expansion and their MO coefficients optimized in the MCSCF procedure. *Deleted* orbitals are the empty analogs of the frozen orbitals. They are empty in all configurations of the CI expansion and their orbital coefficients are not modified by the MCSCF orbital relaxation.

The N electrons that populate the n active orbitals are referred to as the *active electrons* and together with the active orbitals they form the active space, commonly labeled as CAS(N,n).

The full-optimized reaction space (FORS), proposed by Ruedenberg in 1982, [32] is conceptually very similar to the CASSCF. Comprehensive reviews about the initial development of MCSCF procedures are available in the literature [33, 34].

Suffering from an exponential growth [35] of the full-CI expansion with respect to the number of active electrons and orbitals, active spaces beyond CAS(18,18) (coupled to a singlet spin state) are out of reach for conventional approaches (Davidson diagonalization and Direct-CI algorithm for step-vector updates) [7].

Development of methods to circumvent the exponential scaling in MCSCF represents an active area of research. The Stochastic-CASSCF [36] method, the density matrix renormalization group DMRG-SCF [37, 38] method and the *variational*-2RDM [39] method are common approaches to tackle the exponential scaling problem in the MCSCF procedure, in principle without loss of information about the intricate coupling among electronic configurations (in practice for large active spaces containing about 30 electrons and 30 orbitals, these methods provide approximations to the true wave functions and their energies [40]). Other approaches such as the restricted active space, RAS, [41] and the generalized active space [42], reduce the exponential scaling at the cost of removing some of the configurations from the CI expansion. In principle, by using these methods only *deadwood* configurations are removed. In practice, also some important configurations are eliminated from the CI expansion, potentially compromising the accuracy of the procedure.

From an algorithmic point of view current MCSCF procedures can be categorized in two classes: (1) methods based on the direct minimization of the energy, expanded up to second order in some orbital rotations parameters and, (2) methods based on the generalized Brillouin theorem [28, 29, 43–46] (see Section 6.5.8).

Orbitals and CI coefficients can be *coupled* (faster convergence at higher computational costs) or optimized alternately in a *decoupled* procedure (slower convergence and reduced computational costs).

In the initial development of the MCSCF approach, considerable progress was made with respect to full second order procedures (including orbital-CI coupling terms), [34, 47–57] decoupled quasi-second order algorithms (with second order procedures for the orbital optimization step) [30, 31] and, first order methods [28, 29, 43, 44]. In second order MCSCF approaches, gradient and Hessian (first and second derivatives) of the energy with respect to all variational parameters (CI coefficients and MO coefficients) are evaluated exactly. The energy is then approximated by a second order Taylor expansion and a stationary point is found, leading to updated variational parameters. Higher order derivatives may as well be included to increase convergence. This route is rarely undertaken in full, due to the exceedingly high computational demand. Various higher order terms (but not all) are included in the Werner, Meyer, Knowles (WMK) approach [47, 53, 56, 57] and, the advantage of adding these terms is documented. Early MCSCF calculations on relatively large active spaces could not be done without the progress on direct-CI techniques for fast and efficient CI optimizations [3, 4, 6].

6.5.1 The MCSCF Parameterization

In MCSCF theory the wave function is parameterized over the CI coefficients, C_{i0}

$$|0\rangle = \sum_i C_{i0}|\Phi_i\rangle \tag{6.137}$$

and the orbital parameters, $k_{\alpha r}$

$$\varphi_r = \sum_{\alpha} k_{\alpha r} \chi_\alpha \tag{6.138}$$

We assume the orthonormality condition of the wave function

$$\langle \Phi_I | \Phi_J \rangle = \delta_{IJ} \qquad \sum_I C_I^2 = 1 \tag{6.139}$$

and orbitals

$$\langle \varphi_r | \varphi_s \rangle = \delta_{rs} \tag{6.140}$$

The χ_α are approximations to the atomic orbitals (AOs), namely contracted Gaussian functions or Slater-type orbitals. Throughout this section molecular orbitals are assumed to be real functions. The total energy is minimized with respect to both parameters, $E(k, C)$. The variational parameters, $k_{\alpha r}$ and C_{i0}, are introduced in the form of unitary transformations of the wave function. These unitary transformations are realized via exponential representation of the corresponding operators, namely $e^{\hat{R}}$ for the orbital relaxation and, $e^{\hat{S}}$ for the CI optimization. The operators $\hat{R}$ and $\hat{S}$ are anti-Hermitian matrices ($R_{qp} = -R_{pq}^*$). The exponential of an anti-Hermitian matrix is a unitary matrix. When unitary transformations are applied to orthonormalized vectors the resulting vectors are still orthonormal; thus, no orthonormalization procedure is required upon each transformation. Variations of the MCSCF wave function are thus written as

$$|0'\rangle = e^{\hat{R}} e^{\hat{S}} |0\rangle. \tag{6.141}$$

This expression of the MC wave function is also referred to as the *exponential-i-lambda* (EIL) parameterization [49–51]. The operator $\hat{R}$ is defined as

$$\hat{R} = \sum_{pq} R_{pq} \hat{a}_p^\dagger \hat{a}_q. \tag{6.142}$$

The matrix R represents the set of independent orbital rotations parameters. When operating on the wave function, this operator will transform each spin-orbital, φ_s, of the original wave function into a new spin-orbital $\tilde{\varphi}_s$

$$\tilde{\varphi}_s = \sum_r \mathrm{e}^{R_{rs}} \varphi_r \tag{6.143}$$

For real molecular orbitals, the matrix **R** is real and anti-symmetric ($R_{pq} = -R_{qp}$) and the operator takes the compact form

$$\hat{R} = \sum_{p>q} R_{pq} (\hat{E}_{pq} - \hat{E}_{qp}) = \sum_{p>q} R_{pq} \hat{E}_{pq}^- \tag{6.144}$$

The operator $\hat{E}_{pq}^-$ is referred to as the "replacement" operator. The R_{pq} factors are the actual optimization parameters. When stationary condition is reached, the matrix of the optimum R_{pq} elements is used to derive $\mathbf{U} = e^{\mathbf{R}}$ and **U** used to transform the MO coefficients as

$$\tilde{\varphi}_s = \sum_r \varphi_r U_{rs} \tag{6.145}$$

or in matrix form

$$\tilde{\mathbf{X}} = \mathbf{X}\mathbf{U}. \tag{6.146}$$

The operator $\hat{S}$ is defined as

$$\hat{S} = \sum_{K \neq 0} S_{K0} (|K\rangle\langle 0| - |0\rangle\langle K|) \tag{6.147}$$

where $|K\rangle$ represent the complement states orthogonal to the target state $|0\rangle$.

Eq. (6.141) can be thought as first applying the $e^{\hat{S}}$ transformation to $|0\rangle$ and then apply the orbital transformation to the transformed $|\tilde{0}\rangle$ state. In principle, Eq. (6.141) could also be written as

$$|0'\rangle = e^{\hat{S}} e^{\hat{R}} |0\rangle \tag{6.148}$$

with orbital and configuration operators in switched order. This form leads to the unnecessary complication of evaluating the overlap between untransformed and transformed orbitals and for this reason is not considered in practice.

6.5.2 The MCSCF Gradient and Hessian

With the optimization parameters (orbital and CI coefficients) arranged in a column vector, $\mathbf{p}$, we make a Taylor expansion of the energy, $E = E(p)$, around the reference state

$$E(\mathbf{p}) = E(0) + \sum_i \left(\frac{\partial E}{\partial p_i}\right)_0 p_i + \frac{1}{2}\sum_{i,j} p_i \left(\frac{\partial^2 E}{\partial p_i \partial p_j}\right)_0 p_j + \ldots \tag{6.149}$$

The methods used to solve this equation can be divided into two main classes, depending on the rate of convergence. First order methods are based only on the calculation of the energy and its first derivative (first order expansion). Second order methods are based on the calculation of the energy and its first and second derivatives (second order expansion). Higher order expansions are of rather little practical importance due to the exceedingly large computational costs associated to them. However, including higher order derivative terms in an approximate way has been demonstrated important to improve convergence of the second order Taylor expansion [56, 57]. In the local region (that is when the trial CI and orbital parameters are not too far from a stationary point) a second order Taylor expansion ensures a quadratically converging optimization (the energy change decreases quadratically at each iteration). However, in non-local regions second order approximations may not be optimal and second order algorithms require some improvements to guarantee convergence [47, 53, 56, 57].

The Super-CI method, discussed in Section 6.5.8, is a special first order procedure, which has been proven to have near-second order convergence in practical application. The method is not based on direct minimization of the energy expression expanded up to a certain order. Instead, it is better classified as a Brillouin theorem driven approach.

Defining the gradient vector, $\mathbf{E^{(1)}}$, and the Hessian matrix (second derivative), $\mathbf{E^{(2)}}$, with elements

$$E_i^{(1)} = \left(\frac{\partial E}{\partial p_i}\right)_0 \tag{6.150}$$

$$E_{ij}^{(2)} = \left(\frac{\partial^2 E}{\partial p_i \partial p_j}\right)_0 \tag{6.151}$$

one obtains to the second order

$$E(\mathbf{p}) = E(0) + \mathbf{E^{(1)}}^{\dagger}\mathbf{p} + \frac{1}{2}\mathbf{p}^{\dagger}\mathbf{E^{(2)}}\mathbf{p}. \tag{6.152}$$

Setting the derivative of $E(\mathbf{p})$ with respect to all variational parameters, $\mathbf{p}$, equal to zero, a system of inhomogeneous linear equations is obtained

$$\mathbf{E^{(2)}}\mathbf{p} = -\mathbf{E^{(1)}} \tag{6.153}$$

Solving the above equation can be achieved by a sequence of Newton–Raphson iterations, where $\mathbf{E^{(1)}}$ and $\mathbf{E^{(2)}}$ are computed for a given trial wave function, Eq. (6.153) is then solved and, a new vector $\mathbf{p}$ is obtained. This new vector is used to redefine the new point of expansion, rebuild $\mathbf{E^{(1)}}$ and $\mathbf{E^{(2)}}$ and return to Eq. (6.153).

6.5.3 One-Step and Two-Step Procedures

Orbital coefficients and CI coefficients can be separated from the collective column vector, **p**, and the gradient and Hessian written in the following block form

$$\mathbf{E}^{(1)} = \begin{pmatrix} {}^{c}\mathbf{E}^{(1)} \\ {}^{o}\mathbf{E}^{(1)} \end{pmatrix} \tag{6.154}$$

$$\mathbf{E}^{(2)} = \begin{pmatrix} {}^{cc}\mathbf{E}^{(2)} & {}^{co}\mathbf{E}^{(2)} \\ {}^{oc}\mathbf{E}^{(2)} & {}^{oo}\mathbf{E}^{(2)} \end{pmatrix} \tag{6.155}$$

Superscripts *o* and *c* indicate the derivatives with respect to the orbital coefficients and the CI expansion coefficients, respectively. The second derivatives are of three types: orbital–orbital type (*oo*), configuration–configuration type (*cc*), and CI–orbital coupling type (*co*). The block form of the Newton–Raphson equations is then obtained as

$$\begin{pmatrix} {}^{cc}\mathbf{E}^{(2)} & {}^{co}\mathbf{E}^{(2)} \\ {}^{oc}\mathbf{E}^{(2)} & {}^{oo}\mathbf{E}^{(2)} \end{pmatrix} \begin{pmatrix} \mathbf{S} \\ \mathbf{R} \end{pmatrix} = - \begin{pmatrix} {}^{c}\mathbf{E}^{(1)} \\ {}^{o}\mathbf{E}^{(1)} \end{pmatrix} \tag{6.156}$$

The direct solution of Eq. (6.156) is called a *one-step* procedure as **S** and **R** vectors are updated simultaneously. The solution of this equation is not very practical due to the size of the ${}^{cc}\mathbf{E}^{(2)}$ block that scales as M^2, with M of the size of the CI vector. Also, the evaluation of the CI-orbital coupling terms are not easy to evaluate and store. In large MCSCF calculations the explicit construction and storage of the Hessian matrix is not feasible. It is therefore advisable to look for simplifications that avoid some of the difficult computations of the full second order procedure.

Orbital parameters can be separated from the CI parameters by solving for **S** first. From the first row of Eq. (6.156) one obtains

$$\mathbf{S} = -{}^{cc}\mathbf{E}^{(2)^{-1}}({}^{c}\mathbf{E}^{(1)} + {}^{co}\mathbf{E}^{(2)}\mathbf{R}) \tag{6.157}$$

which can be inserted into the second row to obtain an equation in the only variable **R**

$$({}^{oo}\mathbf{E}^{(2)} - {}^{oc}\mathbf{E}^{(2)cc}\mathbf{E}^{(2)^{-1}co}\mathbf{E}^{(2)})\mathbf{R} = {}^{oc}\mathbf{E}^{(2)cc}\mathbf{E}^{(2)^{-1}c}\mathbf{E}^{(1)} - {}^{o}\mathbf{E}^{(1)} \tag{6.158}$$

For a given set of orbitals a CI optimization may be performed first to obtain the initial CI eigenvectors. Being in a stationary point, it follows that

$${}^{c}\mathbf{E}^{(1)} = 0 \tag{6.159}$$

and Eq. (6.158) is simplified to

$$({}^{oo}\mathbf{E}^{(2)} - {}^{oc}\mathbf{E}^{(2)cc}\mathbf{E}^{(2)^{-1}co}\mathbf{E}^{(2)})\mathbf{R} = {}^{o}\mathbf{E}^{(1)} \tag{6.160}$$

This equation is the basis of the *folded two-step Newton–Raphson procedure*. A transformed set of orbitals may be obtained and a new CI optimization carried out. In the two-step procedure the orbital rotation parameters are always obtained on the basis of CI eigenstates. Equation (6.160) is not computationally more advantageous than Eq. (6.158), given among its terms the inverse of the generally large ${}^{cc}\mathbf{E}^{(2)}$ block. This problem could be circumvented only if the ${}^{cc}\mathbf{E}^{(2)}$ matrix is diagonal, that is by solving the CI problem to all orders. And this is generally prohibitively expensive as for very large CI expansion we are not able to obtain all eigensolutions of the CI problem.

A convenient and successful procedure widely used in modern quantum chemistry software packages is to neglect the coupling terms altogether. This leads to the *uncoupled two-step Newton–Raphson procedure* and the equations to solve become

$${}^{oo}\mathbf{E}^{(2)}\mathbf{R} = -{}^{o}\mathbf{E}^{(1)} \tag{6.161}$$

$${}^{cc}\mathbf{E}^{(2)}\mathbf{S} = -{}^{c}\mathbf{E}^{(1)} \tag{6.162}$$

It should be pointed out that this approach is no longer quadratically convergent as the CI-orbital coupling terms are now missing. Including orbital-CI coupling terms is essential in full second order methods as only then the fast quadratic convergence of the macro-iterations is achieved, and only this justifies the expensive computational costs associated with the two indices AO-MO transformations, required in second order procedures. As discussed in Sections 6.5.6–6.5.7, orbital-CI coupling terms can be included during the micro-iterations of the Newton–Raphson method or other methods [47, 56], to a large extent circumventing the complication of solving the CI problem in all orders.

6.5.4 Augmented Hessian Method

In non-local regions the Hessian, $\mathbf{E^{(2)}}$, is generally not positive-definite (it shows negative eigenvalues), thus, compromising the convergence. *Level shift* techniques can be used to make the Hessian positive-definite by introducing a parameter, k, in Eq. (6.153)

$$(\mathbf{E^{(2)}} + k\mathbf{1})\mathbf{p} = -\mathbf{E^{(1)}} \tag{6.163}$$

The level shift parameter can be chosen to be

$$k = -\lambda\epsilon \tag{6.164}$$

with

$$\epsilon = \lambda(\mathbf{E^{(1)}})^{\dagger}\mathbf{p}. \tag{6.165}$$

This choice of the level shift is motivated by the fact that at stationary point the gradient is zero, ϵ vanishes and Eq. (6.163) reduces to Eq. (6.153). Far from stationarity, $\mathbf{E^{(1)}} \neq 0$, the shift parameter, k, provides a positive shift to the Hessian terms, and by adjusting the numerical parameter, λ, the Hessian can be made positive definite. Additionally, the system of inhomogeneous linear equations can be transformed into an eigenvalue problem

$$\begin{pmatrix} -\epsilon & (\mathbf{E}^{(1)})^{\dagger} \\ \mathbf{E}^{(1)} & (\mathbf{E}^{(2)}/\lambda) - \epsilon \end{pmatrix} \begin{pmatrix} 1/\lambda \\ \mathbf{p} \end{pmatrix} = \mathbf{0} \tag{6.166}$$

From the top row of Eq. (6.166) Eq. (6.165) is easily derived and from the bottom row, Eq. (6.163) emerges. The λ parameter add more flexibility to the shift. For $\lambda = 1$, the level shift method is referred to as the *augmented Hessian method* (AHM) [52].

6.5.5 Matrix form of the First and Second Derivatives in MCSCF

We write the energy as a function of the variational parameters

$$E(\mathbf{R}, \mathbf{S}) = \langle 0|e^{-\hat{S}}e^{-\hat{R}}\hat{H}e^{\hat{R}}e^{\hat{S}}|0\rangle \tag{6.167}$$

Recalling the *Baker–Campbell–Hausdorff* (BCH) expansion

$$e^{-\mathbf{A}}\mathbf{B}e^{\mathbf{A}} = \mathbf{B} + [\mathbf{B}, \mathbf{A}] + \frac{1}{2!}[[\mathbf{B}, \mathbf{A}], \mathbf{A}] + \frac{1}{3!}[[[\mathbf{B}, \mathbf{A}], \mathbf{A}], \mathbf{A}] + \ldots \tag{6.168}$$

The second order expansion of the energy expression becomes

$$E(\mathbf{R}, \mathbf{S}) = \langle 0|\hat{H}|0\rangle + \langle 0|[\hat{H}, (\hat{R} + \hat{S})]|0\rangle + \frac{1}{2}\langle 0|[[\hat{H}, (\hat{R} + \hat{S})], (\hat{R} + \hat{S})]|0\rangle \tag{6.169}$$

The first term is the zeroth order energy, the reference energy. The second term collects first derivatives with respect to the orbital and the CI coefficients, respectively. The third term contains second derivatives with respect to both variational parameters.

Orbital Gradient

Expanding the orbital rotation operator in matrix form (Eq. (6.144)), the following result is obtained for the first derivative with respect to the orbital variations

$$\langle 0|[\hat{H}, \hat{R}]|0\rangle = \sum_{p>q} R_{pq} \langle 0|[\hat{H}, \hat{E}^-_{pq}]|0\rangle \tag{6.170}$$

where the R_{pq} elements are the variational coefficients in the orbital space and,

$${}^{o}E^{(1)}_{pq} = \langle 0|[\hat{H}, \hat{E}^-_{pq}]|0\rangle \tag{6.171}$$

the derivative in the pq direction. At stationary points

$$\langle 0|[\hat{H}, \hat{E}^-_{pq}]|0\rangle = 0 \tag{6.172}$$

This condition is referred to as the *generalized (or extended) Brillouin theorem* (GBT) [43, 44], due to the close resemblance to the Brillouin theorem in Hartree–Fock theory. The meaning of this equation is straightforward: there is no interaction, via the Hamiltonian operator, between the optimized wave function $|0\rangle$ and single excitations from it, $\hat{E}^-_{pq}|0\rangle$. Notice that if p and q are both in the same space (inactive, active or virtual) the equation is identically zero for simple rules of second quantization related to the replacement operator, E^-_{pq} (creation of an electron in a occupied orbital or annihilation of an empty orbital equal zero).

Orbital Hessian

For a quadratically converging orbital optimization the orbital Hessian, ${}^{oo}\mathbf{E}^{(2)}$, is also necessary

$${}^{oo}E^{(2)}_{pq,rs} = \langle 0|[[\hat{H}, \hat{E}^-_{pq}], \hat{E}^-_{rs}]|0\rangle. \tag{6.173}$$

The explicit expressions of these matrix elements in terms of one- and two-electron density matrices and integrals can be derived using the simple commutation relations of one- and two-excitation generators in the language of second quantization [30, 31]. The orbital Hessian elements contain integrals of the type $(pq|ij)$ and $(pi|qj)$ where indices p and q run over all orbitals and i and j over the inactive and active orbitals. These two-electron integrals in MO basis are obtained by a four-index transformation. The transformation is sometimes referred to as "second order" transformation as two indices span the entire orbital space. The operation count for the transformation from AO to MO basis is proportional to $n_{occ}N^4_{tot}$ (n_{occ} is the number of active plus inactive orbitals, N_{tot} is the total number of orbitals) and, represents the most time consuming part in the second order orbital optimization procedure when large basis set are employed.

Without considering reduction by space symmetry constraints, the size of the Hessian matrix is $(n_{occ}N_{tot})^2$. For example, if a calculation is performed with $n_{occ} = 150$ and $N_{tot} = 1000$, ~180 GB of memory would be required to store the Hessian elements.

In Section 6.5.6 a commonly used quadratically converging procedure will be discussed in detail, and in Section 6.5.8 the first order Super-CI method will be presented to reduce the computational cost of second order procedures.

CI Gradient

Recalling the definition of the exponential operator for the CI expansion

$$\hat{S} = \sum_{K\neq 0} S_{K0}(|K\rangle\langle 0| - |0\rangle\langle K|) \tag{6.174}$$

the first derivative with respect to the CI parameters is obtained

$$\begin{aligned}\langle 0|[\hat{H},\hat{S}]|0\rangle &= \sum_{K\neq 0} S_{K0}\langle 0|[\hat{H},(|K\rangle\langle 0| - |0\rangle\langle K|)]|0\rangle \\ &= \sum_{K\neq 0} S_{K0}(\langle 0|\hat{H}|K\rangle + \langle K|\hat{H}|0\rangle)\end{aligned} \tag{6.175}$$

In this equation S_{K0} are the variational coefficients and the term in parenthesis the first derivatives. For real wave functions, elements $\langle 0|\hat{H}|K\rangle$ are symmetric therefore the derivative reduces to

$${}^{c}E_{K}^{(1)} = 2\langle 0|\hat{H}|K\rangle \tag{6.176}$$

When a stationary point is obtained the first derivative is zero and the optimized state will not interact, via the Hamiltonian operator, with any state of the orthogonal complement space. This statement is equivalent to saying that the target state $|0\rangle$ is a solution to the secular problem

$$(\mathbf{H} - E\mathbf{1})\mathbf{C} = 0 \tag{6.177}$$

where $\mathbf{H}$ is the Hamiltonian matrix, $\mathbf{C}$ the eigenvectors for the target states and E the corresponding eigenvalues.

CI Hessian

The CI Hessian is expressed as

$${}^{cc}E_{K,L}^{(2)} = \frac{1}{2}\langle 0|[[\hat{H},\hat{S}_K],\hat{S}_L]|0\rangle = 2\langle K|\hat{H} - E^0|L\rangle \tag{6.178}$$

The $\langle K|\hat{H}|L\rangle$ terms are the coupling matrix elements of the MCSCF Hamiltonian over the complement space orthogonal to $|0\rangle$. Their values are accessible only if the CI eigenvalue problem is solved exactly at each MCSCF iteration for all eigenstates. In practice, the CI space is too large to perform a full diagonalization and as a consequence elements $\langle K|\hat{H}|L\rangle$ are, in general, not available. Cheaper methods have been recommended to replace the exact Hessian with approximated ones. In modern chemistry software packages the Davidson approach is the quasi-Newton method of choice [1]. The Davidson procedure can be described as a diagonal approximation of the Hessian in Eq. (6.162). This approximation simplifies the evaluation and inversion of the Hessian matrix during the CI vector update step. The method is presented in more detail in Section 6.4.1.

Although an exact second order optimization in the orbital and CI coefficients is possible, it is in practice rarely done due to the prohibitively large computational cost of evaluating and processing Hessian matrix elements. For practical calculations quasi-second order procedures are adopted for which full knowledge of the Hessian is not required. When a stationary point is reached the gradient is zero. If available, Hessian terms would contain information about the curvature of the hypersurface. When the eigenvalues of the Hessian are all positive a local minimum is reached. Mixed positive and negative values identify saddle points. In two-step and quasi-second order procedures, Hessian terms are not available to be used as indicators of stationarity.

Explicit forms of Gradient and Hessian terms

Gradient and Hessian terms can be further expanded in terms of MO integrals and one- and two-reduced density matrices by inserting the explicit expression of the Hamiltonian operator into Eqs. (6.171), (6.173), (6.176) and (6.178).

By inserting the explicit expression for the Hamiltonian operator, Eq. (1.18) from Chapter 1, into Eq. (6.171) the following expression for the orbital gradient is obtained

$$
\begin{aligned}
{}^{o}E^{(1)}_{xy} &= \langle 0|[\hat{H}, \hat{E}^{-}_{pq}]|0\rangle \\
&= \sum_{pq} h_{pq}\langle 0|[\hat{E}_{pq}, \hat{E}^{-}_{xy}]|0\rangle + \frac{1}{2}\sum_{pqrs}(pq|rs)\langle 0|[\hat{e}_{pqrs}, \hat{E}^{-}_{xy}]|0\rangle.
\end{aligned} \tag{6.179}
$$

Recalling the commutation properties of the one- and two-excitation operators, the definition of one- and two-electron density matrices, the eight-fold permutational symmetry among two-electron integrals and, the four-fold symmetry of the two-electron density matrices we arrive to the following expression for the orbital gradient

$$
{}^{o}E^{(1)}_{xy} = \langle 0|[\hat{H}, \hat{E}^{-}_{xy}]|0\rangle = 2(F_{xy} - F_{yx}) \tag{6.180}
$$

where we have introduced the *generalized Fock matrix* [31]

$$
F_{xy} = \sum_{q} D_{xq} h_{qy} + \sum_{qrs} \Gamma_{xqrs}(yq|rs). \tag{6.181}
$$

Since Γ_{pqrs} terms are in the energy expression traced with the integrals, they are commonly further symmetrized

$$
\Gamma^{s}_{ij,kl} = \frac{1}{2}(\Gamma_{ij,kl} + \Gamma_{ji,kl}). \tag{6.182}
$$

without loss of generality. It follows the eight-fold symmetry for Γ^{s}, in analogy to the two-electron integrals. This additional symmetry can already be utilized in Eq. (6.181).

The generalized Fock matrix is in general not Hermitian (non-symmetric for real functions), and the difference $2(F_{xy} - F_{yx})$ equals the value of the orbital gradient at the point of expansion. It is symmetric only at stationary points, where the gradient is zero. Thus, *a symmetric generalized Fock matrix is a sufficient condition for a stationary wave function with respect to orbital rotations.* This statement is an alternative definition of the generalized Brillouin theorem [43, 44].

The evaluation of the generalized Fock matrix can be simplified if the separation of the orbital spaces in inactive (indices i, j, k, l), active (indices t, u, v, x) and virtual (indices a, b) is taken into account. The elements of the one- and two- electron density matrices vanish if one or more of the MO indices are virtual. Simplifications apply also for density matrices with one inactive index. The result of these simplifications can be summarized as

$$
F_{ip} = 2(F^{I}_{pi} + F^{A}_{pi}) \tag{6.183}
$$

$$
F_{tp} = \sum_{u} F^{I}_{pu} D_{tu} + \sum_{uvx} \Gamma_{tuvx}(pu|vx) \tag{6.184}
$$

$$
F_{ap} = 0 \tag{6.185}
$$

where we have introduced the inactive, $\mathbf{F}^{I}$, and active, $\mathbf{F}^{A}$, Fock matrices with elements

$$
F^{I}_{pq} = h_{pq} + \sum_{i}(2(pq|ii) - (pi|iq)) \tag{6.186}
$$

$$
F^{A}_{pq} = \sum_{tu} D_{tu}\left((pq|tu) - \frac{1}{2}(pu|tq)\right) \tag{6.187}
$$

The index p has been used as generic label. Inactive and active Fock matrices, $\mathbf{F}^{I}$ and $\mathbf{F}^{A}$, are always symmetric, in contrast to the generalized Fock matrix that is symmetric only at stationary points. Additionally, matrices $\mathbf{F}^{I}$ and $\mathbf{F}^{A}$ can easily be computed in AO basis at costs equivalent to Hartree–Fock theory, then transformed into MO basis when needed.

With these equations in hands the operation count and therefore the computational costs can be estimated. For the non-separated generalized Fock matrix (Eq. (6.181)) there are $n_{occ}N_{tot}$ non-zero elements ($n_{occ} = n_{in} + n_{act}$). This follows from the fact that variable y in Eq. (6.181) spans the entire orbital space, while variable x spans only the internal orbital space, as D_{xq} and Γ_{xqrs} vanish for x in the virtual space. The evaluation of each of those elements requires n_{occ} multiplications for the one-electron part and n_{occ}^3 for the two-electron part. Therefore, the operation count for this form of the generalized Fock matrix is, to the leading terms, $N_{tot}n_{occ}^4$. The construction of this matrix is thus a fifth order process.

If we use the form expressed in terms of inactive and active Fock matrix the operation count is much smaller. For the $\mathbf{F}^I$ matrix the total operation count is $N_{tot}^2 n_{in}$ (we have included in the count also those elements with first index virtual, even though they are not needed for the evaluation of the generalized Fock matrix blocks). For the active Fock matrix the total operation count is $N_{tot}^2 n_{act}^2$. The operation count for the term containing the two-body density matrix is $N_{tot}n_{act}^4$. The total of these operations has to be compared to the operation count of the non-separated expression, $N_{tot}n_{occ}^4$, showing that the separation of the generalized Fock matrix in blocks has a clear advantage. It has to be mentioned that in this discussion only the operation count for the construction of the generalized Fock matrix has been analyzed. The AO to MO integral transformation has been omitted. This transformation is expensive (without special techniques it scales as N_{tot}^5) and, usually dominates the computation.

6.5.6 Quadratically Converging Method with Optimal Convergence

Although the AHM or other level shift techniques can improve convergence, they do not completely prevent convergence to be slow, when starting from non-local regions. The cause of this instability in non-local regions has been identified by Werner and Meyer [47, 58] and, it is related to the orthogonality condition over the orbital variations.

For a given set of molecular orbitals and CI coefficients the energy is written as

$$E_0 = \sum_{ij} h_{ij} D_{ij} + \frac{1}{2} \sum_{ijkl} \Gamma_{ijkl} (ij|kl) \tag{6.188}$$

or

$$E_0 = \sum_{ij} \langle i|\hat{h}|j\rangle D_{ij} + \frac{1}{2} \sum_{ijkl} \Gamma_{ijkl} \langle i|\hat{J}_{kl}|j\rangle \tag{6.189}$$

where the generalized Coulomb operator, $\hat{J}_{kl}$ is defined as

$$\langle i|\hat{J}_{kl}|j\rangle = J_{ij}^{kl} = (ij|kl). \tag{6.190}$$

Notice that in the energy expression, Eq. (6.189) labels refer only to internal orbitals, as for external orbitals density matrices $\mathbf{D}$ and $\mathbf{\Gamma}$ vanish and do not provide a contribution to the energy of the MCSCF wave function. If the energy is expanded in the transformed orbitals

$$\tilde{\varphi}_s = \sum_r \varphi_r U_{rs} \tag{6.191}$$

its expression becomes a fourth order function of $\mathbf{U}$ as

$$(ij|kl)(U) = \sum_{pqrs} U_{pi} U_{qj} (pq|rs) U_{rk} U_{sl}, \tag{6.192}$$

and of infinite order of $\mathbf{R}$ ($\mathbf{U} = e^{\hat{\mathbf{R}}} = 1 + \mathbf{R} + \frac{1}{2}\mathbf{R^2} + \ldots$). A second order expansion of the energy in $\mathbf{R}$ is not able to describe correctly the periodic energy variations over the orbital rotations, leading to a slow (or wrong) convergence behavior.

Werner and Meyer [47] and Werner and Knowles [56] realized that convergence of second order procedures can be improved by expanding the orthogonal transformation matrix, $\mathbf{U}$ as

$$\mathbf{U} = \mathbf{1} + \mathbf{T} \tag{6.193}$$

where the transformation matrix, $\mathbf{T}$, can be described as a perturbation of the starting orbitals. The transformed orbitals are then written as

$$\tilde{\varphi}_s = \varphi_s + \sum_r \varphi_r T_{rs} \tag{6.194}$$

The energy expression is thus expanded to the second order in $\mathbf{T}$ (instead of $\mathbf{R}$), and $\mathbf{E}(\mathbf{T^2})$ is minimized, instead of $\mathbf{E}(\mathbf{R^2})$. The second order expansion in $\mathbf{T}$ contains terms of infinite order in $\mathbf{R}$, as opposed to $\mathbf{E}(\mathbf{R^2})$. The explicit energy expression to the second order in $\mathbf{T}$ is simply derived by inserting Eq. (6.194) into Eq. (6.189) for each of the four orbital indices, i, j, k and, l

$$\mathbf{E}(\mathbf{T^2}) = E_0 + 2\sum_{ijw} T_{wi}\langle w|\hat{F}_{ij}|j\rangle + \sum_{ijwx} T_{wi}\langle w|\hat{G}_{ij}|x\rangle T_{xj}, \tag{6.195}$$

where $\hat{F}_{ij}$ and $\hat{G}_{ij}$ are defined as

$$\hat{F}_{ij} = \hat{h}D_{ij} + \sum_{kl} \Gamma^s_{ij,kl}\hat{J}_{kl} \tag{6.196}$$

$$\hat{G}_{ij} = \hat{F}_{ij} + 2\sum_{kl} \Gamma^s_{ik,jl}\hat{K}_{kl} \tag{6.197}$$

with the generalized *exchange* operator $\hat{K}_{kl}$ defined as

$$\langle r|\hat{K}_{kl}|s\rangle = K^{kl}_{rs} = (rk|ls). \tag{6.198}$$

The $\Gamma^s_{ij,kl}$ elements are the *symmetric* contraction of the two-electron density matrices

$$\Gamma^s_{ij,kl} = \frac{1}{2}(\Gamma_{ij,kl} + \Gamma_{ji,kl}). \tag{6.199}$$

Using the expression of the generalized Fock matrix (see Eq. (6.181)), Eq. (6.195) can be rewritten as

$$\mathbf{E}(\mathbf{T^2}) = E_0 + 2\sum_{iw} T_{wi}F_{wi} + \sum_{ijwx} T_{wi}G^{ij}_{wx}T_{xj} \tag{6.200}$$

where also the $\langle w|\hat{G}_{ij}|x\rangle = G^{ij}_{wx}$ equivalence has been used. As D_{pq} and Γ_{pqrs} elements vanish if one or more of the MO indices are virtual, the matrices F_{ij} and G_{ij} can be calculated from the set of operators J_{ij} and K_{ij} with indices i and j spanning only the inactive and active space (also referred as *internal* space). Thus in Eq. (6.195) labels i and j refer to internal orbitals and, w and x to generic orbitals.

The T_{pq} parameters are obtained by solving the following equation

$$\frac{\partial}{\partial T_{pq}}\left(\mathbf{E}(\mathbf{T^2}) - \sum_{ij}\varepsilon_{ij}[(\mathbf{U}^\dagger\mathbf{U})_{ij} - \delta_{ij}]\right) = 0 \tag{6.201}$$

The term multiplied by the Lagrangian multipliers, ε_{ij}, accounts for the orthogonality. Inserting Eq. (6.195) into Eq. (6.201) the following set of nonlinear equations is obtained

$$2\sum_j \langle p|F_{qj}|j\rangle + 2\sum_{jk}\langle p|G_{qj}|k\rangle T_{kj} - 2(\mathbf{U}\varepsilon)_{pq} = 0. \tag{6.202}$$

and defining

$$B_{pq} = 2\sum_j \langle p|F_{qj}|j\rangle + 2\sum_{jk}\langle p|G_{qj}|k\rangle T_{kj} \tag{6.203}$$

Eq. (6.202) can be written in the more compact matrix form

$$\mathbf{U}^\dagger\mathbf{B} - \mathbf{B}^\dagger\mathbf{U} = 0. \tag{6.204}$$

Analogously to $\hat{F}_{ij}$ and $\hat{G}_{ij}$ terms B_{pq} depend on D_{qj} and Γ_{qjkl} and, thus, they vanish whenever q refers to a virtual orbital and, only B_{pq} terms are required with index q in the primary space. It is important to notice that $\mathbf{B}$ depends on $\mathbf{T}$, and the system of equations (6.204) needs to be solved iteratively. These internal iterations are often referred to as *micro-iterations*, to be distinguished from the *macro-iterations* in which a new set of Coulomb and exchange integrals (via a four-index transformation) are evaluated and, a new CI eigenvalue problem solved. In each micro-iteration the updated $\mathbf{B}(\mathbf{T})$ matrix is obtained via a one-index transformation of the one- and two-electron integrals:

$$\tilde{h}_{rj} = (\mathbf{hU})_{rj} \tag{6.205}$$

$$\tilde{J}^{kl}_{rj} = (\mathbf{J}^{kl}\mathbf{U})_{rj} \tag{6.206}$$

$$\tilde{K}^{kl}_{rj} = (\mathbf{K}^{kl}\mathbf{T})_{rj} \tag{6.207}$$

This transformation scales as $\frac{3}{2}N^2n^3$, and is cheap with respect to the four-index transformation needed at each macro-iteration. The second order orbital optimization procedure can thus be summarized as:

- Calculate F_{ij} and G_{ij} (Eqs. (6.196) and (6.197)) from the list of MO integrals, h_{ij}, J_{ijkl} and K_{ijkl} and density matrices.
- Calculate the $\mathbf{B}$ matrix, Eq. (6.203).
- Solve Eq. (6.204) in order to obtain $\mathbf{U}$. As $\mathbf{B}$ depends on $\mathbf{T} = \mathbf{U} - 1$, this equation is solved iteratively (*micro-iteration*).
- Rotate the initial orbitals $\tilde{\mathbf{X}} = \mathbf{XU}$.
- A new *macro-iteration* is started by computing new MO integrals and solving a new CI eigenvalue problem in the new MO basis, $\tilde{\mathbf{X}}$.

At this point we may discuss the computational effort for the algorithm given above. The $\langle p|\hat{G}_{rs}|q\rangle$ matrices are linear combinations of the $\langle p|\hat{J}_{rs}|q\rangle$ and $\langle p|\hat{K}_{rs}|q\rangle$ MO integrals. If r and s refer to active orbitals the total operation count for the $\mathbf{G}$ block is $n^4_{act}N^2_{tot}$. If r refer to active and s to inactive the operation count is $n^2_{act}n_{in}N^2_{tot}$. The operation count for r and s both in the inactive space is only quadratic in the number of inactive orbitals. With the $\mathbf{G}$ matrix elements computed and stored, the evaluation of the $\mathbf{B}$ matrix at each micro-iteration requires $n^2_{occ}N^2_{tot}$ operations. This step is cheaper than the evaluation of the generalized Coulomb and exchange terms; however, it still has a non-negligible impact to the total computational cost.

The evaluation and transformation of the two-electron integrals J_{ij} and K_{ij} is referred to as "second order" transformation as MOs with two indices in the entire orbital space have to be built. This step has a total operation count of about

$$\frac{3}{8}N^4_{tot}n_{occ} + \frac{35}{12}N^3_{tot}n^2_{occ} + \frac{11}{3}N^3_{tot}n_{occ} \tag{6.208}$$

and represents the rate determining step when large basis sets are employed. It should be noted that current chemistry software packages can utilize density fitting (DF) [59] or Cholesky decomposition (CD) [60, 61] approximations, which reduce the operation count to $N_{tot}^3 n_{occ}^2$.

Some approximations have been suggested [47], to reduce the computational costs associated to the inactive orbitals, that reduce the number of operators J_{ij}, K_{ij}, F_{ij} and G_{ij} required. By these approximations the quadratic convergence is lost. It has been recommended not to use first order approximations for core orbitals within the framework of quadratically converging procedures, as they would spoil convergence [56]. Instead a commonly used strategy is to *freeze* energetically low lying core orbitals to the shape determined by the preceding Hartree–Fock approach, here assumed to be the method that generates the starting orbitals for the subsequent MCSCF calculation [56, 62]. This approach certainly guarantees a stable second order convergence. However, it introduces a severe constraint to the MCSCF optimization. Some orbital rotations in the MCSCF procedure, involving core, semi-core or low-lying valence orbitals, are induced by the modified valence mean-field generated by the multi-configurational wave function and having a frozen orbital space certainly compromises the MCSCF relaxation. It is recommended not to have a frozen space for MCSCF optimizations, or, when unavoidable to limit it only to core orbitals energetically well separated from the valence region, namely 1*s* shells for first-row main group atoms and, 1*s*2*s*2*p* shells for first-row transition metal atoms. When dealing with large molecules (long chains would be a better attribute), where only a small portion shows multi-configurational character, frozen orbitals can be chosen in some localized basis rather than in the canonical Hartree–Fock orbital basis. In this specific case, the latter orbital sets are not well suited for the purpose of freezing core orbitals since, they appear rather delocalized and extend from regions of space in proximity of the multi-configurational site and far out to the tail of the system. Much better results are obtained if the HF orbitals are first localized allowing a more obvious separation of the orbitals that could be involved in the MCSCF relaxation (the ones in close proximity of the multi-configurational site) from the orbitals that would only marginally be affected by the MCSCF orbital relaxation. Then the latter orbitals are chosen to be frozen [63]. Alternatively, the Super-CI approach, discussed in Section 6.5.8, is to be considered as a robust alternative, that circumvents the computational limitations associated to second order procedures.

6.5.7 Orbital-CI Coupling Terms

In the two-step decoupled Newton–Raphson approach, where the CI-orbital coupling Hessian blocks are neglected, the second order convergence behavior is in general lost, making any effort to keep a quadratically converging behavior for the orbital part in vain. Quadratically converging procedures are relevant only if the orbital-CI coupling terms are considered in the MCSCF procedure. Energy-based second order MCSCF procedures allow for the orbital-CI coupling in a straightforward way [33, 58, 64]. In this section, the method proposed by Werner and Knowles will be presented [56, 58], that avoids the evaluation of density matrices derivatives and includes the coupling terms via a set of nonlinear equations.

When evaluating the one-index-transformed one- and two-electron integrals (Eqs. 6.205–6.207) also the second half transformation is performed and the following integrals evaluated:

$$h_{ij}^{(2)} = (\mathbf{U}^\dagger \mathbf{h} \mathbf{U})_{ij} \tag{6.209}$$

$$\begin{aligned}(ij|kl)^{(2)} = &-(ij|kl) + (\mathbf{U}^\dagger \mathbf{J}^{kl} \mathbf{U})_{ij} + (\mathbf{U}^\dagger \mathbf{J}^{ij} \mathbf{U})_{kl} \\ &+ (1+\tau_{ij})(1+\tau_{kl})(\mathbf{T}^\dagger \mathbf{K}^{ik} \mathbf{U})_{jl}\end{aligned} \tag{6.210}$$

where τ_{ij} exchanges labels i and j.

These integrals are used to compute an updated Hamiltonian

$$\hat{H}^{(2)} = \sum_{ij} h_{ij}^{(2)} \hat{E}_{ij} + \sum_{ijkl} (ij|kl)^{(2)} \hat{e}_{ijkl}. \tag{6.211}$$

With the updated Hamiltonian, $\hat{H}^{(2)}$, a CI-step is performed (to be distinguished from the CI optimization performed at the beginning of each macro-iteration) to obtain a coupled-corrected CI vector. In the original development of this method the update was performed by the Davidson procedure [56]. The perturbatively updated CI vector is used to build new density matrices, D and Γ and together with the one-index-transformed integrals, Eqs. (6.205–6.207), in a new micro-iteration, in order to obtain a new **B** matrix. It has been found that the perturbative CI update does not have to be performed at each new micro-iteration. Instead it has been suggested that it be performed every 3–5 micro-iterations. This minimized the number of CI updates per macro-iteration, a factor that is certainly important in CI-problem dominated MCSCF optimizations.

6.5.8 Super-CI for the Orbital Optimization

The Super-CI method has been developed to circumvent the limitations in the orbital optimization step of second order procedures. The method was first developed by Grein [45] and Ruedenberg [46] and is based on the Generalized Brillouin theorem (GBT, Eq. (6.172)) [43, 44], which represents the necessary and sufficient condition for a local minimum in the multi-configurational hypersurface. In the context of the uncoupled two-step procedure, we assume that every MCSCF iteration starts by solving the CI problem and as a consequence the optimized ground state wave function, $|0\rangle$, does not interact with the complement space.

We define a Super-CI wave function that contains the ground state wave function, $|0\rangle$, and all possible singly excited states

$$|p \to q\rangle = (\hat{E}_{pq} - \hat{E}_{qp})|0\rangle. \tag{6.212}$$

Here the effect of the excitation operators $(\hat{E}_{pq} - \hat{E}_{qp})$ on each configuration of $|0\rangle$ is that of creating two new configurations in which the orbital of index q is replaced by p and vice versa. The singly excited states, $|p \to q\rangle$ are referred to as *Brillouin states*, because their interaction with $|0\rangle$ via the Hamiltonian operator generates the Brillouin terms (Eq. (6.172)). The Super-CI wave function, $|SCI\rangle$, is therefore defined as

$$|SCI\rangle = |0\rangle + \sum_{p>q} \chi_{pq} |p \to q\rangle. \tag{6.213}$$

The method involves solving the corresponding secular equations to obtain the Super-CI coefficients, χ_{pq}. These coefficients are subsequently used in matrix exponential form to define a unitary transformation of the orbitals in $|0\rangle$ (or some approximation to it [65]). Far from the stationary points the wave function is improved by adding all single excitation contributions to $|0\rangle$. Coefficients χ_{pq} will decrease as the SCF optimization proceeds and at convergence all coefficients will vanish, implying that no more rotations are needed and $|0\rangle$ will reveal the variationally optimized wave function. This condition is equivalent to the Brillouin condition (Eq. (6.172)). Super-CI states are not orthonormalized to each other, in fact

$$\langle 0|\hat{E}_{pq}^{-}\hat{E}_{rs}^{-}|0\rangle \neq \delta_{pq,rs}. \tag{6.214}$$

However, they are orthogonal to the reference state

$$\langle 0|\hat{E}_{pq}^{-}|0\rangle = 0. \tag{6.215}$$

In the particular case of CAS wave functions, only indices p and q belonging to different subspaces have to be considered, as rotations between two orbitals in the same subspace are *redundant*. Thus, only inactive–active, active–virtual and inactive–virtual rotations are considered when generating the $|p \to q\rangle$ states

$$|i \to a\rangle = \hat{E}_{ai}|0\rangle \tag{6.216}$$

$$|t \to a\rangle = \hat{E}_{at}|0\rangle \tag{6.217}$$

$$|i \to t\rangle = \hat{E}_{ti}|0\rangle \tag{6.218}$$

where label i refers to inactive, a to virtual, and t to active orbitals. Normalization factors can be introduced at this stage, but have been neglected here for simplicity. Notice that the Hermitian terms ($\hat{E}_{ia}|0\rangle$, $\hat{E}_{ta}|0\rangle$ and $\hat{E}_{it}|0\rangle$) always vanish, and the Brillouin condition reduces to

$$\langle 0|\hat{H}\hat{E}_{ai}|0\rangle = 0 \tag{6.219}$$

$$\langle 0|\hat{H}\hat{E}_{at}|0\rangle = 0 \tag{6.220}$$

$$\langle 0|\hat{H}\hat{E}_{ti}|0\rangle = 0. \tag{6.221}$$

These equations represent the necessary and sufficient condition for optimized orbitals at each CASSCF macro-iteration. In practice, the Super-CI Hamiltonian matrix is built and diagonalized. Hamiltonian matrix elements between the reference state and the Brillouin states, $\langle 0|\hat{H}|p \to q\rangle$, as well as coupling terms between Brillouin states, $\langle s \leftarrow r|\hat{H}|p \to q\rangle$, are required when solving the Super-CI eigenvalue problem. If solved exactly, the Super-CI secular problem would be more costly than the Newton–Raphson method as it implies the evaluation of up to third order density matrix elements, Q_{vxyztu}, with indices in the active space [30]. However, it is not necessary to use the exact expressions for the $\langle s \leftarrow r|\hat{H}|p \to q\rangle$ terms. Approximations are possible as long as the Brillouin condition is fulfilled.

Based on the work of Hinze [66], Roos and co-workers [28, 29] proposed an approximated generalized-Fock-matrix-based Super-CI approach that avoids the evaluation of third order density matrices [29]. Recalling that the generalized Fock-matrix elements are expressed as

$$F_{xy} = \sum_{q} D_{xq} h_{qy} + \sum_{qrs} \Gamma_{xqrs}(yq|rs). \tag{6.222}$$

and that the orbital gradient $\langle 0|[\hat{H}, \hat{E}^{-}_{pq}]|0\rangle$ can be expressed in terms of the generalized Fock matrix (Eq. (6.180)), the Brillouin terms can be expressed as

$$\langle 0|\hat{H}\hat{E}_{ai}|0\rangle = F_{ai} \tag{6.223}$$

$$\langle 0|\hat{H}\hat{E}_{at}|0\rangle = F_{at} \tag{6.224}$$

$$\langle 0|\hat{H}\hat{E}_{ti}|0\rangle = F_{ti} - F_{it}. \tag{6.225}$$

When calculating matrix elements between singly excited states, $\langle s \leftarrow r|\hat{H}|p \to q\rangle$, terms of the type F_{na} appear, where index, a, refers to secondary orbitals. In this case the generalized Fock-matrix elements vanish (due to the vanishing D_{qa} and Γ_{qrsa} elements). A different expression of the Fock matrix is thus used that leads to non-vanishing coupling terms between Brillouin states. A model one-particle Hamiltonian of the form

$$\hat{H}' = \sum_{pq} \hat{E}_{pq} F'_{pq} \tag{6.226}$$

is introduced and used only for the matrix elements coupling singly excited states, $\langle s \leftarrow r|\hat{H}'|p \rightarrow q\rangle$. $F'_{pq} = F_{pq}$ when p is in the primary space (inactive and active orbitals), otherwise

$$F'_{ab} = F^{ss}_{ab} = h_{ab} + \sum_{pq} D_{pq}\left[(ab|pq) - \frac{1}{2}(ap|bq)\right]. \tag{6.227}$$

The (ss) superscript refers to the fact that this matrix expression is used only for the *secondary subspace*. One important computational aspect arising from the choice of the one-electron model Hamiltonian operator, is that matrix elements do not depend on "second order" MO integral transformations (the two index transformations). Several Super-CI Hamiltonian matrix elements are evaluated directly from a list of atomic one- and two-electron integrals (similar to the closed-shell HF method). A small additional effort is required to evaluate the two-electron MO integrals of the type $(qv|xy)$ where q runs over the entire orbital space while v, x and y are limited to the active space. The transformation of these integrals from AO to MO basis requires about $n_{act}N^4/2$ operations. The second order transformation needed to form $(ab|xy)$ (two general indices) in quadratic procedures requires around $n_{occ}N^4$ operations. For cases with a large number of inactive orbitals the advantage of using the first order procedure is obvious. One- and two-body reduced density matrices are also required in the Super-CI approach presented here (in the exact Super-CI approach also three-body density matrices are required). Two-body density matrices Γ_{pqrs} simplify to first order density matrices if any of the four indices refers to the inactive orbital space. The remaining part contains indices that run only over the active space. None of the two-body density matrix blocks represents a limiting factor. The asymptotic convergence of the Super-CI approach is only first order, and it slows down in proximity of the stationary point. Convergence can be improved by using a quasi-Newton update approach [41].

6.5.9 Redundancy of Active Orbital Rotations

In MCSCF theory some orbital rotations do not lead to energy changes upto first order, and may even interfere with the optimization procedure, introducing singularities in the Hessian and slowing down the iterative procedure that leads to minimization. These rotations are referred to as *redundant*. It is important to identify and eliminate such redundancies. Ways to identify and eliminate redundancies have been suggested in the literature [67]. In CASSCF, given the simple partitioning of the orbital space into inactive, active and virtual orbitals, the identification of redundancies is rather trivial. We recall the orbital rotation operator, introduced in Section 6.5.1

$$\hat{R} = \sum_{p>q} R_{pq}(\hat{E}_{pq} - \hat{E}_{qp}) = \sum_{p>q} R_{pq}\hat{E}^-_{pq} \tag{6.228}$$

where the generic indices p and q independently span the inactive, active and virtual spaces, without restrictions. Rotations within inactive orbitals (p and q inactive) or within virtual (p and q virtual) are redundant as any $\hat{E}^-_{ij}$ or $\hat{E}^-_{ab}$ operators acting on a CAS wave function with i,j inactive and a,b virtual indices equals to zero. This follows from the property of the $\hat{E}_{pq}$ operator that creating electrons in occupied orbitals or annihilating electrons from empty orbitals yields zero. Active–active rotations are also redundant as the effect of any of these orbital rotations is already accounted for by the *complete* CI expansion in the active space that certainly contains configurations of the type $\hat{E}_{tu}|0\rangle$. These easy-to-identify redundancies are set to zero (no intra-space rotations are performed) in order to avoid convergence difficulties.

Rotations that mix different orbital spaces (inactive–active, inactive–virtual, active–virtual) are non-redundant, and are the only ones that need to be addressed in CASSCF orbital optimizations. For more general MCSCF procedures (where the configurational space is not *complete*) the

identification of redundant orbital rotations is more difficult. For instance in RAS or GAS wave functions, where inter-space excitations (say from RAS1 to RAS2) are not *complete*, rotations of orbitals belonging to different spaces are not completely redundant as due to the truncation in the excitation level not all orbital rotation parameters are represented by the CI parameters. As a consequence general MCSCF procedures (including RAS and GAS wave functions) have slower convergence when compared to CASSCF.

6.6 Restricted and Generalized Active Space Wave Functions

In the previous section some theoretical and algorithmic details of MCSCF procedures were introduced, with the main focus on the CASSCF method. In the CASSCF approach a full-CI expansion is built within the active space. The main drawback of the complete expansion in the active space is that the number of configurations scales exponentially with the size of the active space and, with conventional CI eigensolving techniques, active spaces larger than 18 electrons in 18 orbitals are in practice out of reach. The development of large-scale parallel computing [68] and, methods for tackling large CI eigenvalue problems, such as the SplitGAS approach [69, 70], full configuration interaction quantum Monte Carlo (FCIQMC) [36, 40, 71–76], density matrix renormalization group (DMRG) [37, 77–82], the *restricted active space* (RAS) [7, 41, 83, 84] and the *Generalized Active Space* (GAS) approaches [42, 70, 85] have made larger active space problems solvable within MCSCF. The RAS and the GAS approaches are presented in this section. An overview of the FCIQMC algorithm will be given in the next section while the DMRG technique will be presented in Chapter 7.

CI Hamiltonian matrices and eigenvectors are generally extremely sparse. A limited list of electronic configurations completely characterize the electronic wave function of nearly all chemical systems, while most of the configurations do not contribute or contribute only marginally. These *ineffective* configurations (also referred to as *deadwood* configurations) represent up to 99% of the FCI space. Thus, in principle the computational cost of multi-configurational procedures could be greatly reduced with little to no loss in accuracy, by imposing constraints on the active space that remove deadwood configurations. This procedure clearly raises the question of how to identify the unnecessary configurations and selectively choose the effective ones. Several schemes have been devised, based on the partitioning of the active space into several subspaces, and apply restrictions on their electron occupation. Depending on which and how restrictions are chosen, different types of wave functions are obtained. Known procedures are the generalized valence bond (GVB) method [86], constrained CASSCF (C-CASSCF) [87], quasi-CASSCF (Q-CASSCF) [88], Restricted CI (RCI) [89, 90], *occupation restricted multiple active space* (ORMAS) SCF [91], Restricted Active Space (RAS) SCF [7, 41, 83, 84] and, the Generalized Active Space (GAS) SCF [42, 70, 85] approach.

RAS wave functions are obtained by dividing the active orbitals into three subsets, commonly referred to as RAS1, RAS2 and RAS3. The CI expansion is then defined by three parameters: the total number of active electrons in the three RAS spaces, the maximum number of holes in RAS1 $n_{\text{hole},1}$, and the maximum number of particles in RAS3, $n_{\text{elec},3}$. Generally, near doubly occupied orbitals are chosen to be in RAS1, near empty orbitals in RAS3, and the strongly correlated orbitals in RAS2. Compared to CAS wave functions constructed from the same number of active orbitals and electrons, in RAS, configurations that have occupations in RAS1 or RAS3 smaller or larger than the user-specific constraints, are not included in the RAS-CI expansion. RAS wave functions can also be viewed as an extension of CAS wave functions. If no orbitals are chosen for the RAS1 and RAS3 spaces, the RAS scheme reduces to the CAS approach.

The generalized active space (GAS) concept, firstly proposed by Olsen [7, 92], is a further generalization of the RAS concept. Instead of three spaces, in GAS an arbitrary number of active spaces can in principle be chosen. The GAS wave function is completely defined by the following parameters: the number of GAS subspaces, n_{gas}, the list of orbitals belonging to each GAS space and, *accumulated* minimum and maximum electron occupation numbers for the GAS spaces, $n_{\min}(igas)$ and $n_{\max}(igas)$. The ith entry contains the minimum and maximum allowed number of electrons in the first i GAS spaces combined. As a result, $n_{\min}(n_{\text{gas}})$ and $n_{\max}(n_{\text{gas}})$ should equal the total number of active electrons, n_{active}. An example of the GAS approach and the concept of accumulated occupation numbers is given in Table 6.9. In this example the first orbital space, GAS1, contains two to four electrons; the first two spaces together contain four electrons in total; the first three spaces contain from four to five electrons; and finally the four spaces together contain five electrons.

There are two different types of GAS sub-spaces, *disconnected* and *connected*. GAS spaces are referred to as *disconnected* when no inter-space excitations are permitted and, are referred to as *connected* if inter-space excitations are allowed. In the same GAS calculation both connected and disconnected spaces can occur.

It is easy to demonstrate that CAS and RAS wave functions are special cases of GAS wave functions. Regardless of the chosen ansatz, good choices of active spaces still rely on a general understanding of the electronic structure of the system investigated.

For RAS and GAS wave functions, the total number of determinants is reduced because of the constraints imposed on the distribution of electrons among the active orbitals and the exponential scaling partially circumvented. As an example we consider a RAS active space consisting of N electrons and n orbitals, with n_1, n_2, n_3 orbitals in RAS1, RAS2 and RAS3 and, $N_{\alpha 1}$, $N_{\alpha 2}, N_{\alpha 3}$, $N_{\beta 1}$, $N_{\beta 2}, N_{\beta 3}$ alpha and beta electrons distributed in the RAS1, RAS2 and RAS3 spaces of the reference determinant, respectively. Also we choose a maximum of h holes in RAS1 and p electrons in RAS3. The total number of determinants is then given by

$$
\begin{aligned}
N_{SD}(n, N_\alpha, N_\beta) = & \\
\sum_{\substack{N_{\alpha 1}, N_{\beta 1} = n_1 - h \\ N_{\alpha 1} + N_{\beta 1} \geq 2n_1 - h}}^{n_1} & \sum_{\substack{N_{\alpha 3}, N_{\beta 3} = 0 \\ N_{\alpha 3} + N_{\beta 3} \leq p}}^{p} \binom{n_1}{N_{\alpha 1}} \binom{n_1}{N_{\beta 1}} \binom{n_2}{N_\alpha - N_{\alpha 1} - N_{\alpha 3}} \\
& \times \binom{n_2}{N_\beta - N_{\beta 1} - N_{\beta 3}} \binom{n_3}{N_{\alpha 3}} \binom{n_3}{N_{\beta 3}}
\end{aligned}
\tag{6.229}
$$

With an empty RAS2 space, and up to triple excitations from RAS1 to RAS3 (sdt), a total of 40 electrons can be correlated in 40 orbitals. If quadruple excitations are also included (sdtq), the active space size limit is reached with a total of $\sim$ 36 orbitals. The size of the RAS wave function grows rapidly with the number of orbitals in RAS2. For a RAS2 space with 12 orbitals and 12 electrons (in the reference determinant), and considering single and double (sd) RAS1/RAS3 excitations the active space can contain at most 20 orbitals. If triples are also considered (sdt) the maximum number of total active orbitals is reduced to 16. Similar considerations apply to the scaling of GAS wave

Table 6.9 An example of GAS occupation number restriction.

	GAS1	GAS2	GAS3	GAS4
$n_{\min}$	2	4	4	5
$n_{\max}$	4	4	5	5

functions except that due to the flexibility of the GAS approach more GAS spaces and more restrictions to the orbital occupations can be chosen.

6.6.1 GUGA Applied to CAS, RAS and GAS Wave Functions

As for CAS wave functions, CSFs of RAS and GAS wave functions can be represented by GUGA diagrams. In Figure 6.14 a RAS GUGA diagram has been constructed, by eliminating arcs and nodes from the corresponding CAS GUGA diagram (see example in Section 6.3.1 for comparison). The CAS DRT has been listed in Table 6.7 and, Figure 6.9 shows the unconstrained CAS GUGA diagram. The CAS space consists of an active space of five electrons in six orbitals, coupled to a spin $S = \frac{1}{2}$. In RAS the six orbitals are divided into three RAS spaces with two, three and one orbitals, respectively. In RAS1 at most two holes are permitted; and in RAS3 at most one particle is allowed. These restrictions imply that in the GUGA diagram the occupation number at the level of orbital two can only be in the range (2–4), and the occupation number at level of orbital five only (4–5). Hence, vertices 5, 27 and 28 (and all the ones in between) in Figure 6.9 are removed in order to satisfy the RAS constraints and, the GUGA diagram reduces to the one shown in Figure 6.14. The RAS CI space only consists of 106 CSFs, instead of the 210 CSFs of the unconstrained CAS CI expansion. In the RAS GUGA diagram (Figure 6.14), the left edge of the diagram is not altered with respect to the CAS GUGA diagram, and only the vertices at the right hand side are removed with respect to the CAS GUGA diagram. This feature is related to the *closure* property for RAS wave

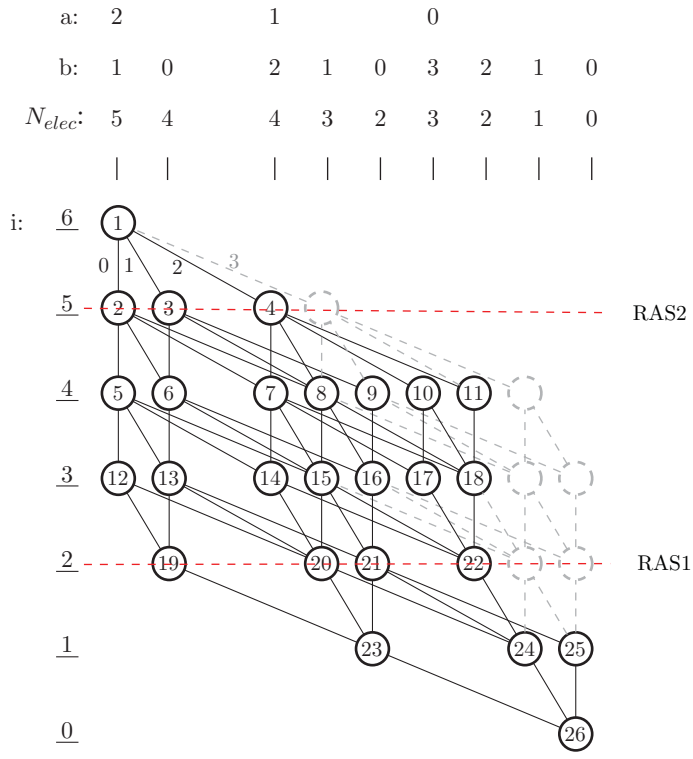


Figure 6.14 Graph representing the DRT of a RAS wave function. The *a* and *b* values indicating the ordering in the graph are shown at the top, and the level index *i* is indicated on the left. The different step-values, *d*, connecting the nodes are shown for the node 1 at the top.

functions, that has some consequences when density matrices and other properties of truncated wave functions are computed (see below) and it represents a distinctive difference with the more general GAS wave functions.

In RAS, for a given active space (RAS1 + RAS2 + RAS3), electron excitations can be divided into *allowed* and *forbidden* excitations according to the RAS occupation constraints. An excitation that leads to a forbidden configuration is also referred to as an *out-of-space excitation*. Out-of-space configurations are part of the unconstrained CAS CI expansion but do not belong to the RAS CI expansion. Since the RAS CI expansion can always be considered as a subspace of the CAS CI expansion, built from the same active space, active–active excitations will always exist that take a certain configuration out of the RAS CI space, but still inside the CAS CI space.

It will now be demonstrated that within the RAS approach, out-of-space excitations never contribute to the corresponding density matrix elements. To show this property we recall the definition of the one- and two-body density matrices

$$D_{tu} = \langle 0|\hat{E}_{tu}|0\rangle = \sum_{IJ} C_I C_J \langle I|\hat{E}_{tu}|J\rangle, \tag{6.230}$$

$$\begin{aligned}\Gamma_{tuvw} &= \langle 0|\hat{e}_{tuvw}|0\rangle = \sum_{IJ} C_I C_J \langle I|\hat{e}_{tuvw}|J\rangle = \\ &= \sum_{IJ} C_I C_J \langle I|\hat{E}_{tu}\hat{E}_{vw}|J\rangle - \sum_{IJ} C_I C_J \langle I|\hat{E}_{tw}|J\rangle \delta_{uv},\end{aligned} \tag{6.231}$$

According to Eq. (6.230), if a single excitation $\hat{E}_{tu}$ excites one electron of a $|J\rangle$ configuration, from RAS1 or into RAS3, so that the resulting configuration does not fulfill the occupation restrictions of the chosen RAS, then this resulting CSF (or SD) will not overlap with any of the $\langle I|$ configurations and the excitation will not contribute to the one-body density matrix element, D_{tu}.

According to Eq. (6.231), for two-body density matrices, double excitations have to be considered. In principle, if an out-of-space excitation $\hat{E}_{vw}$ takes one electron out of RAS1, then $\hat{E}_{tu}$ could still take it back into RAS1 leading to an allowed configuration that overlaps with $\langle I|$ and, thus, provides a contribution to the density matrix element, Γ_{tuvw}. Therefore, in this case an *auxiliary* space would be necessary, of the size of the CAS CI expansion that stores information for the intermediate excitations. However, for RAS wave functions the contribution from the out-of-space excitations to the two-body matrix elements can be removed by using the following permutation relation

$$\hat{E}_{tu}\hat{E}_{vw} = \hat{E}_{vw}\hat{E}_{tu} + \delta_{uv}\hat{E}_{tw} - \delta_{tw}\hat{E}_{vu}, \tag{6.232}$$

defining the ordering of the single excitations $\hat{E}_{tu}$

$$(tu) = (t-1)N + u, \tag{6.233}$$

with N being the number of active orbitals, and considering only elements with $(tu) \geq (vw)$. Under these circumstances if an out-of-space excitation $\hat{E}_{vw}$ takes one electron out of RAS1, then $\hat{E}_{tu}$ is not able to take the electron back into RAS1 given that $(tu) \geq (vw)$. The resulting $\hat{E}_{tu}\hat{E}_{vw}|J\rangle$ configuration would not have any overlap with any $\langle I|$ configuration of the RAS CI expansion, and no contribution to the two-body density matrix element Γ_{tuvw} obtained. This strategy emerges from the *closure* property in RAS wave functions and, represents an efficient way to make the *auxiliary* space not necessary in RAS expansion and thus, to reduce the overall computational cost in RAS calculations.

Table 6.9 summarizes the occupation constraints of a GAS example that further restricts the dimensionality of the CI problem with respect to the CAS(5,6) wave function. Some similarities

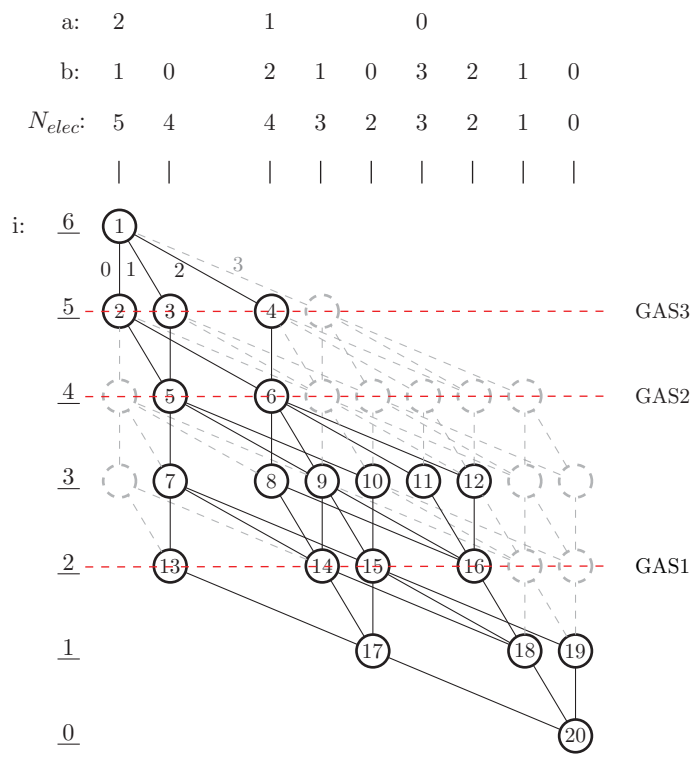


Figure 6.15 Graph representing the DRT of GAS. The *a* and *b* values indicating the ordering in the graph are shown at the top and the level index *i* is indicated at the left. The different step-values *d* connecting the nodes are shown for the node 1 at the top.

can be found between this GAS and the above mentioned RAS constraints, for example GAS1 is equivalent to RAS1. The GAS GUGA diagram is shown in Figure 6.15.

At the level of orbital four the occupation number is restricted to be four, therefore all the vertices and arcs not fulfilling this condition are eliminated. As shown in Figure 6.15, in GAS GUGA diagrams also vertices on the left hand side are eliminated. This property follows from the fact that GAS wave functions are more flexible and more configurations can be eliminated from the CI space. In the above example only 52 CSFs remain in the GAS CI expansion. Meanwhile the property of *closure* discussed above for RAS wave function is lost for GAS wave functions and some complications arise when building the density matrices or using GAS wave functions as reference wave functions for post-GASSCF approaches, such as in GASPT2 [85].

6.6.2 Redundancies in GASSCF Orbital Rotations

As already discussed in Section 6.5.9, active–active orbital rotations are already described by the complete CI expansion of the CAS wave function and are, redundant. Difficulties in the MCSCF optimization procedure may arise, if these redundant rotations are included in the CASSCF orbital relaxation. Also inactive–inactive and virtual–virtual rotations are redundant. These redundant rotations are easy to identify and eliminate. Therefore, in CASSCF procedures only inter-space orbital rotations are kept.

For GASSCF, the identification and elimination of redundant orbital rotations is a bit more complicated as it is dependent on the user choice of the inter-space excitations. A practical choice is to include all inter-space orbital rotations and to eliminate all the intra-space orbital rotations in analogy to CASSCF. This means that GAS1-GAS2, GAS1-GAS3, GAS2-GAS3, and so on, are considered during the GASSCF orbital relaxation. This, however, implies that some of the redundant rotations are included, namely the ones that are represented in the CI space by inter-space electron excitations, when *connected* GAS spaces are considered. Connected GAS sub-spaces cause inter-space orbital rotations to become redundant.

The parameters to define a certain configurational space is not unique in the GAS ansatz. A CAS wave function will be used as an example to show this point. The easiest way to build a CAS wave function, from a GAS perspective, is simply to set $n_{\text{gas}} = 1$ and $n_{\text{min}} = n_{\text{max}} = n_{\text{active}}$. The orbital rotations included in the orbital optimization procedure are all the inter-space rotations, namely the inactive–active, inactive–virtual and, active–virtual. No redundancy in orbital rotations appears. The same configurational space can also be obtained in different ways, for example, by splitting the active space into two spaces and allowing all possible excitations between these two spaces. This form consists of two connected GAS spaces and, although they lead to exactly the same CI wave function as the simple CAS construction, it is obvious that the GAS form contains extra orbital rotations to optimize, namely the GAS1–GAS2 rotations. These rotations are all redundant, as they are already described by the CI expansion. This example shows that generalizations towards more flexible ansatz spaces, within MCSCF, can introduce redundancies, which have to be taken into account to prevent difficulties in the convergence. If only part, but not all, of the excitations among GAS spaces are allowed, it becomes more difficult to identify and easily eliminate the redundant orbital rotations.

6.6.3 MCSCF Molecular Orbitals

CASSCF wave functions are invariant with respect to inactive–inactive, active–active, and virtual–virtual orbital rotations. Therefore, several invariant orbital transformations are possible, which do not mix orbitals of different spaces within each other. Generally, inactive and virtual orbitals are rotated into *canonical* orbitals. This is obtained by block diagonalizing the Hartree–Fock Fock matrix for the inactive and virtual orbitals separately. The eigenvalues are the orbital energies of the MOs, and the eigenvectors are the canonical orbitals. Orbitals of the active space can be represented in a number of ways.

By diagonalizing the one-body density matrix in the active space, the eigenfunctions and eigenvalues lead to the active *natural orbitals* and their occupation numbers, respectively. This procedure is invariant only for CAS wave functions. For RAS and GAS wave functions this procedure is not invariant, due to the inter-space rotations. One can also block-diagonalize the active space density matrix within each RAS (or GAS) space, instead of the whole active space. The obtained orbitals are called *pseudonatural orbitals*. The transformation to pseudonatural orbitals is invariant. For state-averaged MCSCF calculations, several CI wave functions are optimized simultaneously, sharing the same set of MOs. Each state is characterized by its own density matrices, which can be diagonalized separately (or block-diagonalized for RAS and GAS) to obtain one set of natural orbitals (or pseudonatural orbitals) and corresponding occupation numbers per state. These orbitals are commonly used to calculate state-specific properties. During the CASSCF macro-iterations, however, density matrices are weight-averaged, and the weight-averaged energy is minimized with respect to a common set of MOs. The natural orbitals of averaged density matrices are generally also referred to as pseudonatural orbitals.

If the generalization of the Fock matrix for multi-configurational wave functions

$$F_{pq} = h_{pq} + \sum rsD_{rs}\left[(pq|rs) - \frac{1}{2}(pr|qs)\right] \tag{6.234}$$

is diagonalized, *canonical orbitals* are obtained, the eigenvalues representing the corresponding orbital energies. Analogously to the natural orbitals case, this procedure is invariant only for CAS wave functions. For RAS and GAS wave function pseudo-canonical orbitals might be generated, which also represent invariant rotations. These orbitals are generally used for spectral properties and as reference orbitals for subsequent PT2 treatments.

6.6.4 GASSCF Applied to the Gd_2 Molecule

Gd_2 is the diatomic molecule with highest spin multiplicity in its ground state known to date, its ground state being a $^{19}\Sigma_g^-$, corresponding to the following electronic configuration $(\sigma_g^1\sigma_u^1\pi_g^2\pi_u^2\delta_g^2\delta_u^2\phi_g^2\phi_u^2)_{4f}\sigma_g^2\sigma_g^1\pi_u^2\sigma_u^1$. The experimentally determined ground state vibrational constant and dissociation energy are $\omega_e = 138.7 \pm 0.4$ cm^{-1} and $D_e = 2.1 \pm 0.7$ eV respectively [93]. Here we present the CASSCF and GASSCF ground state potential energy curve of Gd_2 [42] and compare them with the experimental values.

A reasonable CAS choice for the entire potential energy curve of Gd_2, is the full valence shell active space, which consists of the 20 valence electrons and 32 molecular orbitals, arising from 4*f*, 5*d*, 6*s* and 6*p* AOs of each Gd atom. This CAS(20,32) generates about 1.4 billion Slater determinants for the $^{19}\Sigma_g^-$ state, currently impossible to tackle with standard methods. The smaller CAS(20,26) is a possible alternative. This active space is obtained by removing the six 6*p* orbitals from the CAS(20,32). This wave function contains 2.1 million determinants. A GAS-2(20,26), analogous to the CAS(20,26), is built by dividing the active space into two *disconnected* GAS spaces: GAS1 contains the 14 4*f* orbitals and, GAS2 the rest. This restriction leaves only 23808 determinants in the wave function. A GAS-5(20,32) with 5 *disconnected* GAS spaces can be utilized to correlate all 32 valence orbitals and, thus, overcome the computational limitations of the CAS(20,32). Of the 32 active orbitals, the 14 4*f* orbitals are separated into four different GAS spaces, based on space symmetry consideration. The remaining six electrons and 18 orbitals form the fifth space. The resulting CI space contains only 138 304 determinants, which is four orders of magnitude smaller than CAS(20,32). The optimized active orbitals of the GAS-5(20,32) and their distribution among the five GAS space is shown in Figure 6.16. The ground state, $^{19}\Sigma_g^-$, potential energy curves for the Gd_2 system at the CAS(20,26), GAS-2(20,26) and GAS-5(20,32) level of theory are presented in Figure 6.17. The curves obtained by the CAS(20,26) and GAS-2(20,26) approaches are almost identical, indicating that mostly deadwood configurations have been removed in the GAS-2(20,26) with respect to the CAS(20,26) and, the accuracy of the MCSCF calculation is preserved. However, the two curves are not smooth throughout the dissociation pathway. This is due to orbital flipping (in and out of the active space) along the dissociation and, the active orbitals in the bonding region and in the dissociation region are different. In the region R=5.20 Å to 5.40 Å, the active orbitals $5d\pi_u$ and $5d\pi_g$ are progressively replaced by the $6p\pi_u$ and $6p\pi_g$ orbitals. The CAS(20,26) and GAS-2(20,26) are not large enough to accommodate all competing correlating orbitals necessary to describe the dissociation path consistently. This behavior is cured in the GAS-5(20,32) approach discussed above, which includes all the relevant orbitals along the dissociation. For this choice of active space orbital-flipping is not observed, orbitals of the same character are correlated along the reaction and, as a result a smooth potential energy curve is obtained.

By fitting the GAS-5(20,32) potential energy curve to a Morse potential, the dissociation energy D_e and vibrational constant ω_e were determined to be 2.1 eV and 140 cm^{-1}, respectively. These values

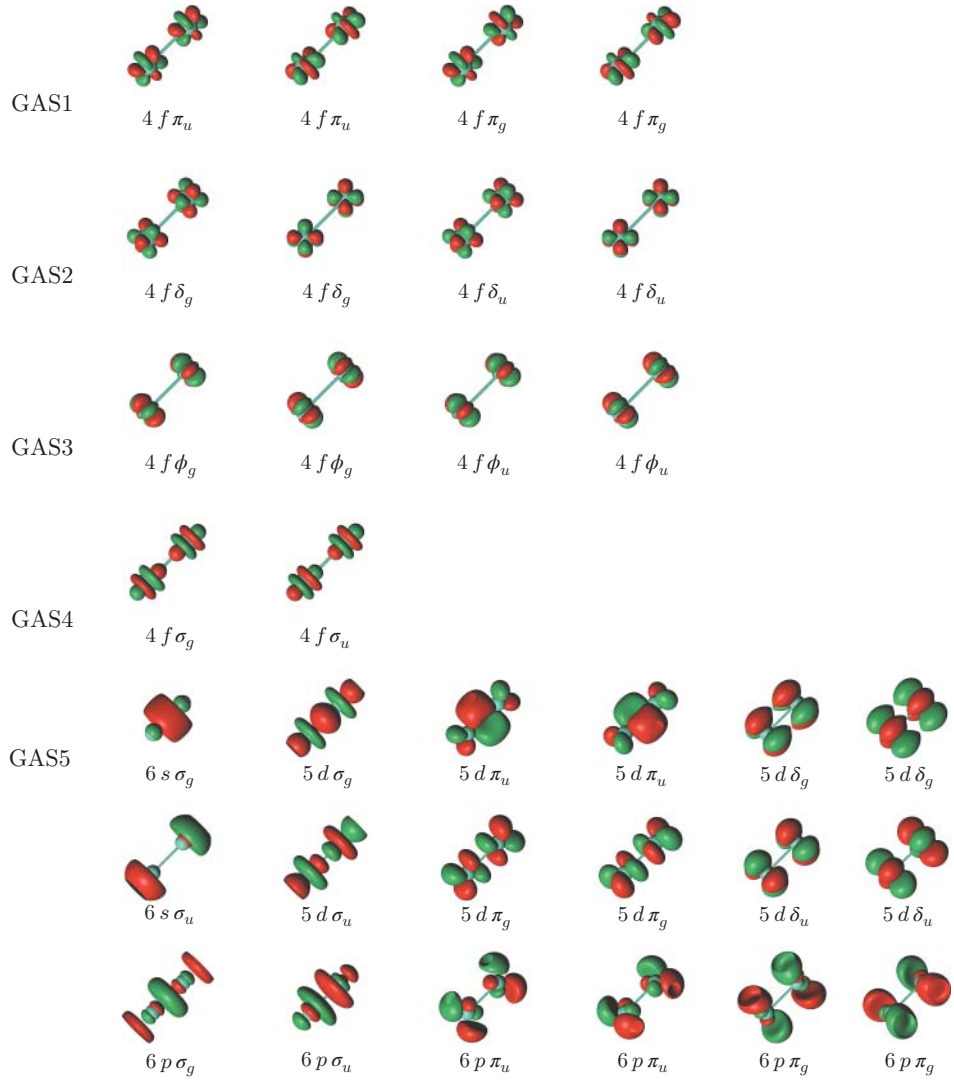


Figure 6.16 Natural orbitals of Gd_2 GAS-5(20,32) at equilibrium bond distance [42].

are in good agreement with the experimental values, demonstrating the strength of the GASSCF method.

6.7 Excited States

An important property of MCSCF procedures is the possibility to simultaneously optimize ground and excited states, while accounting for correlation and relaxation effects, with no bias for any of the computed states. This feature, however, relies on two conditions. First, the CI eigensolver must be able to simultaneously optimize multiple low-lying states. Second, the orthogonality of the resulting states has to be preserved by the MCSCF procedure or, it has to be restored by a posteriori procedures.

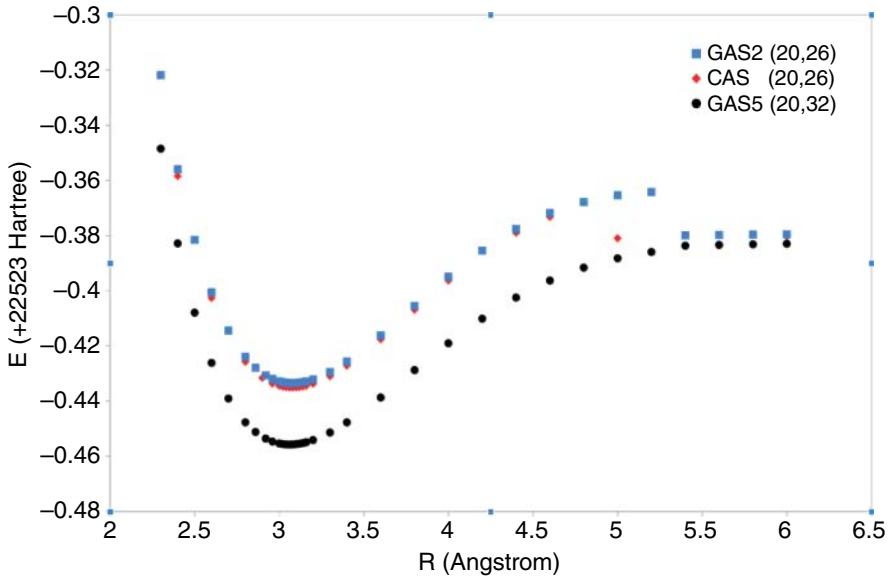


Figure 6.17 Gd_2 $^{19}\Sigma_g^-$ potential-energy curves by using CAS(20,26), GAS-2(20,26) and GAS-5(20,32) [42].

6.7.1 Multi-State CI Solver

If several eigenvalues of a Hermitian matrix H are to be computed, a straightforward approach which can be applied to a wide class of CI solvers is to restrict the search space to the *orthogonal complement* (oc) of a set of already converged eigenvectors. If n eigenvectors $\mathbf{v}_i$ of H are already known, construct the matrix

$$\boldsymbol{U} = [\mathbf{v}_1, \dots, \mathbf{v}_n] \tag{6.235}$$

and then apply the CI solver to the restricted Hamiltonian

$$H|_{\text{oc}} = (\mathbf{1} - UU^\dagger)H(\mathbf{1} - UU^\dagger) \tag{6.236}$$

where $(\mathbf{1} - \boldsymbol{U}\boldsymbol{U}^\dagger)$ projects out the components of the $\mathbf{v}_n$ eigenvectors from the trial wave function during the optimization. Hence, when using $H|_{\text{oc}}$, only orthogonal components will contribute to the target vector. A little caution has to be applied to the circumstances that may lead to trivial solutions of $H|_{\text{oc}}$, namely the ones corresponding to zero eigenvalues. This problem can be circumvented by shifting the Hamiltonian.

By iteratively applying this scheme, excited states can subsequently be generated, finding a state orthogonal to the previous ones in each iteration and thus creating a set of orthonormal eigenvectors. This is typically limited by the precision of the orthogonalization, as numerical inaccuracy leads to an increasing loss of orthogonality of the computed vectors as more and more vectors are added.

Often, however, more efficient schemes can be applied for targeting multiple states. For Krylov-subspace type methods, a straightforward extension of the single-state schemes discussed in Section 6.4.1 can be made by adding multiple states to the Krylov space per iteration, leading to the Block Lanczos Method [94].

This increases the dimension of the ansatz space by up to n per iteration, n being the number of targeted states.

This approach, however, comes with the drawback, that to obtain the nth excited state, all previous states have to be computed too, making it costly or even impossible to obtain higher excited states. An alternative approach to target eigenvalues in the proximity of a given value τ is the family of shift-and-invert methods, which search for the largest eigenvalue of a transformed matrix

$$\boldsymbol{A} = (H - \tau)^{-1} \tag{6.237}$$

For the Davidson method, that can be implemented without the need to explicitly compute the inverse, by using a modified ansatz space that is

$$K' = \text{span}\{(H - \tau)\mathbf{c} | c \in K\} \ . \tag{6.238}$$

See Reference 2 for more details.

6.7.2 State-Specific and State-Averaged MCSCF

Targeting multiple states within one single MCSCF calculation comes with the complication that a straightforward optimization of the orbital basis for each of the targeted states will produce as many orbital sets as the number of optimized states.

In the state-specific CASSCF method, a dedicated CASSCF calculation is performed for each of the targeted states. In each calculation, only the density matrices of the targeted state are used for the orbital optimization step. Typically, even for the state-specific approach, the CI eigensolver requires to compute several CI states in each calculation, namely for the jth state, j states have to be calculated. In addition to the added computational costs (j vectors are stored and processed), in practical applications this approach is plagued by convergence problems.

The state-specific CASSCF approach yields non-orthogonal states and orbitals. A subsequent diagonalization of the Hamiltonian in the basis of the optimized states is thus required.

Another common approach to circumvent the difficulty of different orbital sets for different CI eigenstates is the state-averaged CASSCF technique, which uses weight averaged density matrices

$$\tilde{D} = \sum_{i=1}^{n} w_i D_i \tag{6.239}$$

$$\tilde{\Gamma} = \sum_{i=1}^{n} w_i \Gamma_i, \tag{6.240}$$

to minimize the averaged energy of the target states in a common single set of orbitals. The weight factors, w_i, determine the impact of each single state on the orbital relaxation. Weights can either be chosen equal or tuned to enhance convergence. The resulting orbitals are in general not optimal for any particular state and in general deteriorate when the number of states is enlarged. The advantage of this procedure is that states are expressed in a shared set of orbitals and are therefore orthogonal.

The SCF schemes discussed for ground state MCSCF procedure are unchanged except for the usage of the state-averaged density matrices instead of the state specific density matrices, making the approach straightforward to extend in principle to any number of states.

6.8 Stochastic Multiconfigurational Approaches

In Section 6.1, the computational limitations of MCSCF procedures related to the exponential growth of the wave function with respect to the number of correlated electrons and orbitals has been discussed. These limitations can be substantially reduced by treating the CI eigenvalue problem

stochastically, circumventing the need for storing the full CI vector. Stochastic algorithms to find solutions to the CI eigenvalue problem can be designed in numerous ways, both by projective and variational techniques [95]. With stochastic CI eigensolvers, active spaces of significantly increased size become treatable, with CI vectors containing up to $\sim 10^{18}$ Slater determinants. Within the CASSCF framework, active spaces with up to 40 electrons and 38 orbitals have been reported with unprecedented accuracy [40, 76].

In this section, stochastic CI eigensolvers are discussed using the example of the Full Configuration Interaction Quantum Monte Carlo (FCIQMC) algorithm, outlining the calculation of reduced density matrices from a stochastically sampled wave-function. Its extension to MCSCF procedures is straightforward, relying on the stochastic sampling of the active space one- and two-body density matrices [74] and their usage within approximated uncoupled approaches for the orbital optimization step [36, 75]. Stochastic second-order procedures with CI-orbital coupling as discussed in Section 6.5.6 and 6.5.7 are - to our knowledge - not available to date.

6.8.1 FCIQMC Working Equation

Similar to the deterministic iterative algorithms, in the FCIQMC ground state optimizations, all but the ground state contributions are eliminated from some initial wave function, $|\psi(0)\rangle$. At convergence this procedure will reveal the true ground state wave function. Conversely to the deterministic approaches, mostly based on the solution of the stationary time-independent Schrödinger equation, the FCIQMC approach and many others electronic structure QMC algorithms aim at solving the *imaginary-time Schrödinger equation*

$$i\frac{\partial}{\partial t}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle \xrightarrow{t=-i\tau} -\frac{\partial}{\partial \tau}|\psi(\tau)\rangle = \hat{H}|\psi(\tau)\rangle\,, \tag{6.241}$$

which is simply obtained from the Schrödinger equation by substituting $t = -i\tau$.

The formal solution to the imaginary-time Schrödinger equation is given by

$$|\psi(\tau)\rangle = \mathrm{e}^{-\tau\hat{H}}|\psi(0)\rangle \equiv \hat{U}(\tau)|\psi(0)\rangle. \tag{6.242}$$

An important property of the imaginary-time propagator is that it decays exponentially with τ and, since adding a constant to the Hamiltonian does not affect the eigenvectors, we can assume the energy E_0 of the ground state $|\psi_0\rangle$ of $\hat{H}$, (with $\hat{H}|\psi_0\rangle = E_0|\psi_0\rangle$) to be equal to 0.[4] With $E_0 = 0$ one can easily verify that

$$\lim_{\tau\to\infty} \mathrm{e}^{-\tau\hat{H}}|\psi(\tau)\rangle \to |\psi_0\rangle \tag{6.243}$$

given that

$$\langle\psi(0)|\psi_0\rangle \neq 0\,, \tag{6.244}$$

since all excited state contributions $|\psi_i\rangle$ with $E_i > E_0$ get exponentially suppressed.

The stochastic imaginary time evolution (Eqn. (6.241)) of the ground state wave function in the space of Slater determinants is at the core of the FCIQMC method [71, 72, 98].

To evaluate the imaginary-time evolution, the propagation is carried out repeatedly for small time steps $\Delta\tau$, such that the full propagator $\hat{U}(\Delta\tau)$ can be approximated by the first-order Taylor expansion

$$\hat{U}(\Delta\tau) \approx \hat{U}^{(1)}(\Delta\tau) = \hat{\mathbf{1}} - \Delta\tau(\hat{H} - S(\tau)\hat{\mathbf{1}})\,, \tag{6.245}$$

4 However, we will come back to the practical implementation of this constant shift in the next section.

with an variable energy shift $S(\tau)$. Above, the possibility of adding a constant to the Hamiltonian to shift the ground state energy to 0 was mentioned. This is precisely what $S(\tau)$ is required for. As the ground state energy is unknown in the beginning of the calculation, $S(\tau)$ will be updated iteratively; converging to the correct ground state energy, $\lim_{\tau\to\infty} S(\tau) = E_0$.

The wave function at time $\tau + \Delta\tau$ is obtained from the wave function $|\psi(\tau)\rangle$ at time τ by a single application of $\hat{U}^{(1)}$ as

$$|\psi(\tau + \Delta\tau)\rangle = \hat{U}^{(1)}(\Delta\tau)|\psi(\tau)\rangle \ . \tag{6.246}$$

For a sufficiently small timestep $\Delta\tau$ the approximated propagation converges to the exact ground state, as can be readily seen from expanding $|\psi(0)\rangle$ in an eigenbasis of $\hat{H}$ and applying $\hat{U}^1(\Delta\tau)^n$.

To evaluate $\hat{H}|\psi\rangle$, the FCIQMC method works on a sparsely sampled wave function in Slater determinant space. A sampled wave function

$$|\psi(\tau)\rangle = \sum_i c_i(\tau)|D_i\rangle \ , \tag{6.247}$$

is stored, where $c_i(\tau) \neq 0$ only for a small fraction of the Hilbert space at each iteration; these determinants are referred to as *occupied determinants*. Notably, the wave function used in FCIQMC is **not** normalized, and to profit from the sparse structure of the wave function, a minimal value $c_{\min}$, typically $c_{\min} = 1$, for $|c_i|$ is imposed for the coefficients to be stored, with any coefficient below $c_{\min}$ being stochastically rounded. This is done via a stochastic round function R, which is given by

$$R(x) = \begin{cases} x & \text{if } x \geq c_{\min}, \\ c_{\min} & \text{if } x < c_{\min} \text{ with prob. } \frac{x}{c_{\min}} \\ 0 & \text{else,} \end{cases} \ , \tag{6.248}$$

such that, for example, a coefficient of $c_i(\tau) = 0.6\ c_{\min}$ is rounded up to $c_{\min}$ with a probability of 60% and discarded with a probability of 40%.

Doing so keeps the number of occupied determinants under control, as less important determinants will get neglected more often, removing the need of storing their coefficients. The L_1-norm

$$\|\psi(\tau)\|_1 = \sum_i |c_i(\tau)| \tag{6.249}$$

of $|\psi(\tau)\rangle$ plays the role of a control parameter, as increasing it leads to more coefficients being above $c_{\min}$ and therefore sampling the wave function more precisely. In the limit of infinite L_1 norm, the method becomes exact as all coefficients will be stored.

The striking advantage of a sparse sampling of the wave function, as used in the FCIQMC method is, that the number of occupied determinants in each iteration is typically much lower than the total number of determinants forming the Hilbert space of the problem.

Inserting the representation (6.247) into Eq. (6.246) yields the *working equations* of FCIQMC as

$$\begin{aligned} |\psi(\tau + \Delta\tau)\rangle &= [\hat{\mathbf{1}} - \Delta\tau(\hat{H} - S(\tau)\hat{\mathbf{1}})] \sum_i c_i(\tau)|D_i\rangle \\ &= \sum_i (1 - \Delta\tau(H_{ii} - S(\tau)))c_i(\tau)|D_i\rangle - \Delta\tau \sum_{i,j\neq i} H_{ji}c_i(\tau)|D_j\rangle, \end{aligned} \tag{6.250}$$

where the identity $\hat{\mathbf{1}} = \sum_j |D_j\rangle\langle D_j|$ was inserted.

The evaluation of $\hat{U}^{(1)}(\Delta\tau)|\psi(\tau)\rangle$ is performed in *three steps*, of which the first two calculate $\Delta\tau\hat{H}|\psi(\tau)\rangle$ and the third one returns the propagated wave function. We now carry out these three algorithmic steps of FCIQMC, namely **(1)** *spawning*, **(2)** *death* and **(3)** *annihilation*.

(1) In the spawning step of a FCIQMC iteration, the off-diagonal part of Eq. (6.250) is evaluated by sampling the sum over j stochastically, that is, we randomly pick some matrix elements to apply, neglecting the others while making sure that on average, the correct action of $\hat{H}$ is obtained.

Consider now the contribution $\Delta|\psi\rangle_i$ from a single determinant $|D_i\rangle$, which is

$$\Delta|\psi\rangle_i = -\Delta\tau \sum_{j\neq i} H_{ji} c_i(\tau)|D_j\rangle \ . \tag{6.251}$$

To approximate the sum over j, we first have to decide on how many matrix elements to select. To use the coefficient $c_i(\tau)$ as a number of sampled matrix elements has proven to be a sensible choice, as this approximates the contributions from important determinants better than those from unimportant ones. Since $c_i(\tau)$ is not necessarily an integer, it is stochastically rounded to an integer number M. If for example $c_i(\tau) = 1.5$, M will be chosen to be 1 in 50% of the cases and 2 in the other 50%. The sign of $c_i(\tau)$ has to be kept separately, as it cannot be contained in a (positive) integer. Then, the off-diagonal contribution can be expressed as

$$\Delta|\psi\rangle_i \approx -\mathrm{sgn}(c_i(\tau)) \sum_{k=1}^{M} \sum_{j\neq i} H_{ji}\Delta\tau|D_j\rangle \ . \tag{6.252}$$

For each k, the second sum is now approximated by randomly selecting just one single value for j, which is denoted as j_k, and neglecting all the other terms. We therefor introduce probabilities $p_{\mathrm{gen}}(j|i)$ of selecting determinant $|D_j\rangle$ for a given value of i. As, on average, a determinant $|D_j\rangle$ will then appear $M\ p_{\mathrm{gen}}(j|i)$, whereas it should appear just M times, we need to unbias by dividing by $p_{\mathrm{gen}}(j|i)$. The resulting expression is now

$$\Delta|\psi\rangle_i \approx -\sum_k \mathrm{sgn}(c_i(\tau)) \frac{H_{j_k i}\Delta\tau}{p_{\mathrm{gen}}(j_k|i)}|D_{j_k}\rangle \ , \tag{6.253}$$

For large values of M, this expression becomes more and more accurate, as each j gets picked $M\ p_{\mathrm{gen}}(j|i)$ times on average, eventually yielding the original sum in the limit of $M \to \infty$. $\Delta|\psi\rangle_i$ still includes a potentially large number of determinants, with potentially very small amplitudes. However, as a minimal value c_{min} is imposed on the coefficients, the new contributions to the coefficients of the sampled $|D_{j_k}\rangle$ have to be rounded using the stochastic round R from Eq. (6.248) if their absolute value is below c_{min}. If the outcome of the round is 0, the contribution does not have to be stored. This ensures that the number of accounted determinants does not grow rapidly to unmanageable numbers. The full off-diagonal contribution is then obtained by summing over all determinants with non-zero coefficient. The exact off-diagonal contribution is then replaced with the stochastic approximation

$$\Delta|\psi^{\mathrm{off}}_{\mathrm{approx}}\rangle = -\sum_i \sum_k \mathrm{sgn}(c_i) R\left(\frac{H_{j_k i}\Delta\tau}{p_{\mathrm{gen}}(j_k|i)}\right)|D_{j_k}\rangle. \tag{6.254}$$

The newly created contributions $\Delta|\psi^{\mathrm{off}}_{\mathrm{approx}}\rangle$ will be added to the coefficient of $|D_{j_k}\rangle$ in the next iteration, but are kept separately as *spawns* for now, which will be added to the existing ones later in step (3). This is important as they do not contribute to the diagonal part in step (2).

(2) The death step takes care of the diagonal contribution of Eq. (6.250), which is carried out by multiplying each coefficient $c_i(\tau)$ with $(1 - \Delta\tau(H_{ii} - S(\tau)))$. Again, only those determinants have to be considered, which have non-zero coefficients in the sampled wave function, for all others, the contribution is simply zero. While $\Delta\tau(H_{ii} - S(\tau))c_i(\tau)$ can be obtained with little effort, again, the minimal value of $c_i(\tau)$ has to be guaranteed. Therefore, if $c_i(\tau) < c_{\mathrm{min}}$, $c_i(\tau)$ is stochastically rounded to c_{min}.

In practice, having $\Delta\tau(H_{ii} - S(\tau)) > 1$ can lead to instabilities, as it allows the coefficients to change drastically within one iteration, and should be prevented by choosing a sufficiently small timestep $\Delta\tau$.

(3) In the annihilation step, the new contributions $\Delta|\psi^{\text{off}}_{\text{approx}}\rangle$ to the coefficients from the spawn step are added back to the sampled wave function. We now update the wave function as

$$|\psi(\tau + \Delta\tau)\rangle = \sum_i R([1 - \Delta\tau(H_{ii} - S(\tau))]c_i(\tau))|D_i\rangle + \Delta|\psi^{\text{off}}_{\text{approx}}\rangle \ . \tag{6.255}$$

Obtaining a good value for the variable shift $S(\tau)$ is essential for preventing the norm of the state to diverge or vanish, since Eq. (6.243) assumed the ground state energy of $\hat{H} - S\hat{\mathbf{1}}$ to be 0. Any other average value will inadvertently lead to divergence or vanishing of the norm. This is directly related to the L_1 norm of the state, which dictates the numerical cost of an iteration. In this way, $S(\tau)$ becomes an instrument for controlling the L_1 norm: By lowering $S(\tau)$, we can increase $\|\psi\|_1$ and vice versa. Ideally, $S(\tau)$ should be equal to the ground state energy, but requiring this value as an input defeats the purpose, such that the value of the shift is iteratively updated to maintain a constant $\|\psi\|_1$. A very efficient way of updating $S(\tau)$ is to adjust it every m steps based on the change in $\|\psi\|_1$ during these m steps. As $S(\tau)$ enters the evolution of $\|\psi\|_1$ exponentially, the update should be of logarithmic form. Adding a damping χ to prevent large oscillations leads to the update instruction

$$S(\tau + n\Delta\tau) = S(\tau) - \frac{\chi}{m\Delta\tau} \log\left(\frac{\|\psi(\tau + n\Delta\tau)\|_1}{\|\psi(\tau)\|_1}\right) \ . \tag{6.256}$$

Since any other value of S would lead to either vanishing or diverging norm, in a converged calculation at least the average of $S(\tau)$ has to be the ground state energy, given the sampling of $|\psi(\tau)\rangle$ is accurate. This energy value itself is only of minor practical relevance since it comes with comparably large fluctuations, but it can be used as a consistency check. A more stable estimate for the energy is the projected energy onto a suited reference determinant $|D_0\rangle$, given by

$$E_{\text{proj}} = \frac{\langle D_0|\hat{H}|\psi\rangle}{\langle D_0|\psi\rangle} \ . \tag{6.257}$$

This projected energy is the value of the energy eigenvalue the state $|\psi\rangle$ had if it were an eigenstate of $\hat{H}$. Hence, it is not variational, but becomes exact in the exact limit of FCIQMC.

Because the sampling of the wave function is of stochastic nature, and for any τ, it only represents a "snapshot" of the full ground state wave function, derived quantities like the projected energy carry stochastic noise, creating the need to average them over numerous iterations once convergence is reached.

There are further improvements and modifications of the FCIQMC method, such as the initiator approximation (i-FCIQMC) [72, 73] or the semi-stochastic method [98, 99], which can enhance the potential of the approach and allow for study of substantially larger-sized systems with FCI accuracy.

Especially the initiator approximation is of great importance, as it addresses the sign problem arising within FCIQMC. While the antisymmetry of the wave function is guaranteed by working in Slater determinant space, the CI coefficients themselves can have both positive and negative sign. For determinants with a small coefficient (possibly just c_{min}), that coefficient, by stochastic error, can carry the wrong sign. In the unmodified algorithm, these sign errors quickly propagate.

6.8.2 Multi-Wave Function Approach for Excited States

The FCIQMC method sketched so far can only target the ground state, but an excited state search can be realized by projecting out lower energy states in the spirit of using a restricted Hamiltonian [98], similar to Eq. (6.236) mentioned in Section 6.7.1.

However, projecting out a previously converged state seems like a poor approach, as we will be using only a snapshot of the full wave function. When projecting out a previously obtained state, in the second calculation there is no guarantee that we will not end up with contributions from the ground state which were not part of that snapshot.

A more promising and efficient method is to perform multiple FCIQMC calculations in parallel, evolving n wave functions $|\psi^{(1)}\rangle$ to $|\psi^{(n)}\rangle$ at the same time. Then, in each iteration, the kth wave function is orthogonalized against the first $k-1$ wave functions. In practice, this is realized by using the modified imaginary-time evolution

$$|\psi^{(k)}(\tau+\Delta\tau)\rangle = \hat{P}_k(\tau+\Delta\tau)\hat{U}^{(1)}(\Delta\tau)|\psi^{(k)}(\tau)\rangle\,, \tag{6.258}$$

where we define the instantaneous projector

$$\hat{P}_k(\tau) = \hat{\mathbf{1}} - \sum_{l<k} \frac{|\psi^{(l)}(\tau)\rangle\langle\psi^{(l)}(\tau)|}{\langle\psi^{(l)}(\tau)|\psi^{(l)}(\tau)\rangle}\,, \tag{6.259}$$

which acts as a projection operator on the wave function snapshots $|\psi^{(1)}\rangle \dots |\psi^{(k-1)}\rangle$. The projection is carried out instantaneously, that is, using $\hat{P}_k(\tau+\Delta\tau)$ to obtain $|\psi^{(k)}(\tau+\Delta\tau)\rangle$, which is possible since $\hat{P}_k$ only depends on the instantaneous wave functions $|\psi^{(l)}(\tau+\Delta\tau)\rangle$ with $l<k$, such that the projection can be carried out one after the other starting from $k=1$. The modified evolution can then be seen as performing one iteration with all n wave functions and subsequently orthogonalizing them by removing the components of the new wave function proportional to all $|\psi^{(l)}(\tau+\Delta\tau)\rangle$ with $l<k$. This procedure maintains a total of n orthogonal wave functions. As the first one is not orthogonalized against any other wave function, it behaves just as in the ground state calculation, eventually converging to the ground state. The second wave function then is orthogonal to the ground state, and thus converges to the first excited state, and so on.

6.8.3 Sampling Reduced Density Matrices

To use the FCIQMC method as the CI solver in a CASSCF calculation, we need a way to obtain the one- and two-body reduced density matrices (RDMs)

$$D_{ij} = \langle\psi_0|a_i^\dagger a_j|\psi_0\rangle, \tag{6.260}$$

$$\Gamma_{ijkl} = \langle\psi_0|a_i^\dagger a_j^\dagger a_k a_l|\psi_0\rangle. \tag{6.261}$$

Computing an approximation of the full expectation value with the sampled wave function $|\psi(\tau)\rangle$ as an estimate for $|\psi_0\rangle$, scales quadratically in the number of determinants sampled in $|\psi(\tau)\rangle$, quickly making the full evaluation unfeasible. But recall that any quantity derived from the instantaneous wave function is carrying stochastic noise and therefore needs to be averaged anyway. So adding additional noise by evaluating the expectation value stochastically seems to be tolerable, and reduces the cost of obtaining the RDMs to a minimum.

To understand how the sampling can be performed in an efficient way [74, 99], consider the one-body RDM

$$D_{ij} = \sum_{kl}\langle D_k|a_i^\dagger a_j|D_l\rangle c_k^*(\tau)c_l(\tau)\,. \tag{6.262}$$

Algorithm 3 FCIQMC algorithm

- Pick an initial state
loop
  for all occupied determinants $|D_i\rangle$ **do**
    set $M = \lfloor c_i \rfloor$
    with probability $c_i(\tau) - M$, increase M by 1
    for all $k = 1, \dots, M$ **do**
      - randomly pick a determinant $|D_{j_k}\rangle$ (with $H_{ij_k} \neq 0$
      - Create a spawn of weight $\frac{H_{j_k i}\Delta\tau}{p_{gen}(j_k|i)}$ on $|D_{j_k}\rangle$
      - Round the spawn stochastically to a minimal value if necessary
    end for
    - Multiply $c_i(\tau)$ by $1 - (H_{ii} - S(\tau))$, then roundstochastically if necessary
  end for
  for all determinants $|D_i\rangle$ **do**
    - Sum up $c_i(\tau)$ and all spawns to obtain new $c_i(\tau + \Delta\tau)$
  end for
end loop

Here, for each determinant $|D_l\rangle$ contained in $|\psi(\tau)\rangle$, we need the products of its coefficient $c_l(\tau)$ with the coefficients $c_k(\tau)$ of any single-excitation $|D_k\rangle$ of $|D_l\rangle$. The stochastic evaluation of this sum can be carried out by only taking some of the single excitations, weighted with the probability of picking them. However, this random choice of excitations for each determinant is anyway a part of the core FCIQMC algorithm (3). Therefore, when performing the spawning step, we can easily get a stochastic evaluation of the one-body RDM by logging all spawning attempts to single-excitations and adding the weighted product of the coefficients to the D_{ij}. Doing so yields a stochastic expectation value of D_{ij} for each iteration, and averaging over numerous iterations, after convergence with respect to τ, gives reliable values for the one-body RDM.

Unsurprisingly, the method to evaluate the two-body RDM is analogous, with the only difference being that instead of the spawning attempts to single-excitations, those to double-excitations have to be logged now.

At this point, one might object that the RDMs are *quadratic* in the stochastic quantity $|\psi(\tau)\rangle$ and must therefore be biased, as noise cannot fully average out. And indeed, the quadratic dependence on the coefficients does introduce a bias. If however, the two coefficients $c_k(\tau)$ and $c_l(\tau)$ in Eq. (6.262) come from different, independent samplings of the wave function, this problem would not occur, as they then also carry different, independent noise. Hence, to obtain the RDMs, two FCIQMC calculations have to be performed in parallel, yielding two independent samplings $|\psi^{(a,b)}\rangle$ of the same state. Then, computing estimates for the RDMs as

$$D_{ij} = \langle \psi^{(a)} | a_i^\dagger a_j | \psi^{(b)} \rangle, \tag{6.263}$$

$$\Gamma_{ijkl} = \langle \psi^{(a)} | a_i^\dagger a_j^\dagger a_k a_l | \psi^{(b)} \rangle, \tag{6.264}$$

"unbiases" the expectation values, such that they now average to the correct value.

While the produced stochastic error is controllable and vanishes in the limit of infinite L_1 norm, this sampling also introduces a potential systematic error, which is linked to how the excitations are generated in the FCIQMC spawning step. Consider some determinant $|D_k\rangle$ with a significant coefficient $c_k(\tau)$. Spawning attempts to determinants with a vanishing matrix element will never be successful but those determinants might very well contribute to the RDMs. Consider a single- or double-excitation $|D_l\rangle$ of $|D_k\rangle$ with $\langle D_k|H|D_l\rangle = 0$ but a significant coefficient $c_l(\tau)$,

then the product $c_l^*(\tau)c_k(\tau)$ is never taken into account for the evaluation of the RDMs, even if it is non-negligible.

While in general, we can assume this special case to be an exception, when working with Hartree–Fock orbitals, there is a relevant case which has to be taken care of, and this is the Hartree–Fock determinant. The Hartree–Fock determinant has almost always significant weight and, due to the Brillouin theorem, is not coupled to its single-excitations, even though they often contribute significantly to the RDMs. An efficient and straightforward way to still get the correct contribution to the RDMs from the HF determinant is now to just calculate these deterministically by explicitly taking into account all products of the coefficient of the Hartree–Fock determinant with the sampled single- and double-excitations when computing the RDMs.

Bibliography

1 Davidson, E.R. (1975). The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices. *J. Comput. Phys.* 17 (1): 87–94.

2 Sleijpen, G.L.G. and Van der Vorst, H.A. (1996). A Jacobi-Davidson iteration method for linear eigenvalue problems. *SIAM J. Matrix Anal. Appl.* 17 (2): 401–425.

3 Siegbahn, P.E.M. (1984). A new direct CI method for large CI expansions in a small orbital space. *Chem. Phys. Lett.* 109: 417.

4 Knowles, P.J. and Handy, N.C. (1984). A new determinant-based full configuration interaction method. *Chem. Phys. Lett.* 111: 315.

5 Handy, N.C. (1980). Multi-root configuration-interaction calculations. *Chem. Phys. Lett.* 74 (2): 280–283.

6 Roos, B.O. (1972). New method for large-scale CI calculations. *Chem. Phys. Lett.* 15 (2): 153.

7 Olsen, J., Roos, B.O., Jørgensen, P., and Jensen, H.J.Aa. (1988). Determinant based configuration interaction algorithms for complete and restricted configuration interaction spaces. *J. Chem. Phys.* 89: 2185–2192.

8 Duch, W. (1986). *GRMS or Graphical Representation of Model Spaces*. Berlin: Springer.

9 Shavitt, I. (1978). Matrix element evaluation in the unitary group approach to the electron correlation problem. *Int. J. Quantum Chem.* 14 (S12): 5–32.

10 Shavitt, I. (1993). The history and evolution of Gaussian basis sets. *Isr. J. Chem.* 33: 357.

11 Moshinsky, M. (1968). *Group Theory and the Many-body Problem*. N.Y.: Gordon and Breach.

12 Paldus, J. (1974). Group theoretical approach to the configuration interaction and perturbation theory calculations for atomic and molecular systems. *J. Chem. Phys.* 61 (12): 5321–5330.

13 Shavitt, I. (1977). Graph theoretical concepts for the unitary group approach to the many-electron correlation problem. *Int. J. Quantum Chem.* 12 (S11): 131–148.

14 Cetlin, M.L. and Gel'fand, I.M. (1950). Finite-dimensional representations of the group of unimodular matrices. *Doklady Akad. Nauk SSSR (N.S.)* 71: 825–828.

15 Cetlin, M.L. and Gel'fand, I.M. (1950). Finite-dimensional representations of the group of orthogonal matrices. *Doklady Akad. Nauk SSSR (N.S.)* 71: 1017–1020. *Amer. Math. Soc. Transl.* 64: 116 (1967).

16 Gel'fand, I.M. (1950). The center of an infinitesimal group ring. *Mat. Sb. (N.S.)* 26 (68): 103–112.

17 Louck, J.D. (1970). Recent progress toward a theory of tensor operators in the unitary groups. *Am. J. Phys.* 38: 3–42.

18 Shavitt, I. (1981). The graphical unitary group approach and its application to direct configuration interaction calculations. In: *The Unitary Group for the Evaluation of Electronic Energy Matrix Elements* (ed. J. Hinze), 51–99. Berlin, Heidelberg: Springer Berlin Heidelberg.

19 Paldus, J. and Boyle, M.J. (1980). Unitary group approach to the many-electron correlation problem via graphical methods of spin algebras. *Phys. Scr.* 21: 295.

20 Drake, G.W.F. and Schlesinger, M. (1977). Vector-coupling approach to orbital and spin-dependent tableau matrix elements in the theory of complex spectra. *Phys. Rev. A* 15: 1990–1999.

21 Payne, P.W. (1982). Matrix element factorization in the unitary group approach for configuration interaction calculations. *Int. J. Quantum Chem.* 22 (6): 1085–1152.

22 Jacobi, C.G.J. (1846). Über ein leichtes Verfahren, die in der Theorie der Säcularstörungen vorkommenden Gleichungen numerisch aufzulösen. *Crelle's J.* 30: 51–94.

23 Mises, R.V. and Pollaczek-Geiringer, H. (1929). Praktische Verfahren der Gleichungsauflösung. *ZAMM - J. Appl. Math. Mech. / Zeitschrift für Angewandte Mathematik und Mechanik* 9 (2): 152–164.

24 Lanczos, C. (1950). An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *J. Nat. Bur. Stand.* 45: 225–282.

25 Krylov, A.N. (1931). On the numerical solution of the equation by which in technical questions frequencies of small oscillations of material systems are determined. *Izvestija AN SSSR (News of Academy of Sciences of the USSR), Otdel. mat. i estest. nauk* 7 (4): 491–539.

26 Helgaker, T., Jørgensen, P., and Olsen, J. (2000). *Molecular Electronic Structure Theory*. Chichester, England: Wiley.

27 Davidon, W.C. (1991). Variable metric method for minimization. *SIAM J. Optim.* 1 (1): 1–17.

28 Roos, B.O., Taylor, P.R., and Siegbahn, P.E.M. (1980). A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach. *Chem. Phys.* 48: 157–173.

29 Roos, B.O. (1980). The complete active space SCF method in a Fock-matrix-based super-CI formulation. *Int. J. Quantum Chem.* 18 (S14): 175–189.

30 Siegbahn, P.E.M., Heiberg, A., Roos, B.O., and Levy, B. (1980). A comparison of the super-CI and the Newton-Raphson scheme in the complete active space SCF method. *Phys. Scr.* 21: 323–327.

31 Siegbahn, P.E.M., Almlöf, J., Heiberg, A., and Roos, B.O. (1981). The complete active space SCF (CASSCF) method in a Newton-Raphson formulation with application to the HNO molecule. *J. Chem. Phys.* 74: 2384–2396.

32 Ruedenberg, K., Schmidt, M.W., Gilbert, M.M., and Elbert, S.T. (1982). Are atoms intrinsic to molecular electronic wave functions? I. The FORS model. *Chem. Phys.* 71: 41–49.

33 Shepard, R. (1987). The multiconfiguration self-consistent field method. *Adv. Chem. Phys.* 69: 63.

34 Olsen, J., Yeager, D.L., and Jørgensen, P. (1983). Optimization and characterization of a multiconfigurational self-consistent field (MCSCF) state. *Adv. Chem. Phys.* 54: 1.

35 Aquilante, F., Autschbach, J., Carlson, R.K. et al. (2016). Molcas 8: new capabilities for multiconfigurational quantum chemical calculations across the periodic table. *J. Comput. Chem.* 37 (5): 506–541.

36 Li Manni, G., Smart, S.D., and Alavi, A. (2016). Combining the complete active space self-consistent field method and the full configuration interaction quantum Monte-Carlo within a super-CI framework, with application to challenging metal-porphyrins. *J. Chem. Theory Comput.* 12 (3): 1245–1258.

37 Zgid, D. and Nooijen, M. (2008). The density matrix renormalization group self-consistent field method: orbital optimization with the density matrix renormalization group method in the active space. *J. Chem. Phys.* 128: 144116.

38 Ghosh, D., Hachmann, J., Yanai, T., and Chan, G.K.-L. (2008). Orbital optimization in the density matrix renormalization group, with applications to polyenes and β-carotene. *J. Chem. Phys.* 128: 144117.

39 Fosso-Tande, J., Nguyen, T.-S., Gidofalvi, G., and DePrince, A.E. III (2016). Large-scale variational two-electron reduced-density-matrix-driven complete active space self-consistent field methods. *J. Chem. Theory Comput.* 12: 2260–2271.

40 Li Manni, G., Kats, D., Tew, D.P., and Alavi, A. (2019). Role of valence and semi-core electron correlation on spin gaps in Fe(II)-porphyrins. *J. Chem. Theory Comput.*

41 Malmqvist, P.-Å., Rendell, A., and Roos, B.O. (1990). The restricted active space self-consistent-field method, implemented with a split graph unitary-group approach. *J. Phys. Chem.* 94 (14): 5477–5482.

42 Ma, D., Li Manni, G., and Gagliardi, L. (2011). The generalized active space concept in multiconfigurational self-consistent field methods. *J. Chem. Phys.* 135: 044128.

43 Levy, B. and Berthier, G. (1968). Generalized Brillouin theorem for multiconfigurational SCF theories. *Int. J. Quantum Chem.* 2 (2): 307–319.

44 Levy, B. and Berthier, G. (1968). Generalized Brillouin theorem for multiconfigurational SCF theories. *Int. J. Quantum Chem.* 3 (2): 247–247.

45 Banerjee, A. and Grein, F. (1977). Multiconfiguration wave functions for excited states. Selection of optimal configurations: the $b^1\Sigma^+$ and $d^1\Sigma^+$ states of NH. *J. Chem. Phys.* 66: 1054–1062.

46 Ruedenberg, K., Cheung, L.M., and Elbert, S.T. (1979). MCSCF optimization through combined use of natural orbitals and the Brillouin-Levy-Berthier theorem. *Int. J. Quantum Chem.* 16: 1069–1101.

47 Werner, H.-J. and Meyer, W. (1980). A quadratically convergent multiconfiguration-self-consistent field method with simultaneous optimization of orbitals and CI coefficients. *J. Chem. Phys.* 73: 2342.

48 Shepard, R. and Simons, J. (1980). Multiconfigurational wave function optimization using the unitary group method. *Int. J. Quantum Chem.* 18: 211.

49 Dalgaard, E. and Jørgensen, P. (1978). Optimization of orbitals for multiconfigurational reference states. *J. Chem. Phys.* 69: 3833.

50 Yeager, D.L. and Jørgensen, P. (1979). Convergency studies of second and approximate second order multiconfigurational Hartree-Fock procedures. *J. Chem. Phys.* 71: 755.

51 Dalgaard, E. (1979). A quadratically convergent reference state optimization procedure. *Chem. Phys. Lett.* 65: 559.

52 Lengsfield, B.H. III (1980). General second order MCSCF theory: a density matrix directed algorithm. *J. Chem. Phys.* 73: 382.

53 Werner, H.-J. and Meyer, W. (1981). A quadratically convergent MCSCF method for the simultaneous optimization of several states. *J. Chem. Phys.* 74: 5794.

54 Olsen, J. and Jørgensen, P. (1982). Update methods in multiconfigurational self-consistent field calculations. *J. Chem. Phys.* 77: 6109.

55 Olsen, J., Jørgensen, P., and Yeager, D.L. (1982). Multiconfigurational Hartree-Fock studies of avoided curve crossing using the Newton-Raphson technique. *J. Chem. Phys.* 76: 527.

56 Werner, H.-J. and Knowles, P.J. (1985). A second order multiconfiguration SCF procedure with optimum convergence. *J. Chem. Phys.* 82: 5053.

57 Knowles, P.J. and Werner, H.-J. (1985). An efficient second-order MCSCF method for long configuration expansions. *Chem. Phys. Lett.* 115: 259.

58 Werner, H.-J. (1987). Matrix-formulated direct multiconfiguration self-consistent field and multiconfiguration reference configuration-interaction methods. *Adv. Chem. Phys.* 69: 1.

59 Gyorffy, W., Shiozaki, T., Knizia, G., and Werner, H.-J. (2013). Analytical energy gradients for second-order multireference perturbation theory using density fitting. *J. Chem. Phys.* 138 (10): 104104.

60 Aquilante, F., Malmqvist, P.-Å., Pedersen, T.B. et al. (2008). Cholesky decomposition-based multiconfiguration second-order perturbation theory (CD-CASPT2): application to the spin-state energetics of Co-III(diiminato)(NPh). *J. Chem. Theory Comput.* 4: 694.

61 Aquilante, F., Pedersen, T.B., and Lindh, R. (2007). Low-cost evaluation of the exchange Fock matrix from Cholesky and density fitting representations of the electron repulsion integrals. *J. Chem. Phys.* 126: 194106.

62 Ma, Y., Knecht, S., Keller, S., and Reiher, M. (2017). Second-order self-consistent-field density-matrix renormalization group. *J. Chem. Theory Comput.* 13: 2533–2549.

63 Menezes, F., Kats, D., and Werner, H.-J. (2016). Local complete active space second-order perturbation theory using pair natural orbitals (PNO-CASPT2). *J. Chem. Phys.* 145: 124115.

64 Péter, G., Szalay, T.M., Gidofalvi, G. et al. (2012). Multiconfiguration self-consistent field and multireference configuration interaction methods and applications. *Chem. Rev.* 112 (1): 108–181.

65 Banerjee, A. and Grein, F. (1976). Convergence behavior of some multiconfiguration methods. *Int. J. Quantum Chem.* 10: 123–134.

66 Hinze, J. (1973). MCSCF. I. The multi-configuration self-consistent-field method. *J. Chem. Phys.* 59 (12): 6424–6432.

67 Hoffmann, M.R., Fox, D.J., Gaw, J.F. et al. (1984). Analytic energy second derivatives for general MCSCF wave functions. *J. Chem. Phys.* 80: 2660.

68 Vogiatzis, K.D., Ma, D., Olsen, J. et al. (2017). Pushing configuration-interaction to the limit: towards massively parallel MCSCF calculations. *J. Chem. Phys.* 147 (18): 184111.

69 Li Manni, G., Ma, D., Aquilante, F. et al. (2013). Split-GAS method for strong correlation and the challenging case of Cr_2. *J. Chem. Theory Comput.* 9: 3375–3384.

70 Vogiatzis, K.D., Li Manni, G., Stoneburner, S.J. et al. (2015). Systematic expansion of active spaces beyond the CASSCF limit: a GASSCF/Split-GAS benchmark study. *J. Chem. Theory Comput.* 11: 3010–3021.

71 Booth, G.H., Thom, A.J.W., and Alavi, A. (2009). Fermion Monte Carlo without fixed nodes: A game of life, death and annihilation in Slater determinant space. *J. Chem. Phys.* 131: 054106.

72 Cleland, D., Booth, G.H., and Alavi, A. (2010). Survival of the fittest: accelerating convergence in full configuration-interaction quantum Monte Carlo. *J. Chem. Phys.* 132: 041103.

73 Cleland, D., Booth, G.H., and Alavi, A. (2011). A study of electron affinites using the initiator approach to full configuration interaction quantum Monte Carlo. *J. Chem. Phys.* 134: 024112.

74 Overy, C., Booth, G.H., Blunt, N.S. et al. (2014). Unbiased reduced density matrices and electronic properties from full configuration interaction quantum Monte Carlo. *J. Chem. Phys.* 141: 244117.

75 Thomas, R.E., Sun, Q., Alavi, A., and Booth, G.H. (2015). Stochastic multiconfigurational self-consistent field theory. *J. Chem. Theory Comput.* 11 (11): 5316–5325.

76 Li Manni, G. and Alavi, A. (2018). Understanding the mechanism stabilizing intermediate spin states in Fe(II)-porphyrin. *J. Phys. Chem. A* 122: 4935–4947.

77 White, S.R. (1992). Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.* 69: 2863–2866.

78 Schollwöck, U. (2005). The density-matrix renormalization group. *Rev. Mod. Phys.* 77: 259–315.

79 Chan, G.K.-L. and Sharma, S. (2011). The density matrix renormalization group in quantum chemistry. *Annu. Rev. Phys. Chem.* 62 (1): 465–481.

80 Marti, K.H. and Reiher, M. (2010). The density matrix renormalization group algorithm in quantum chemistry. *Z. Phys. Chem.* 224 (3): 583–599.

81 Olivares-Amaya, R., Hu, W., Nakatani, N. et al. (2015). The ab-initio density matrix renormalization group in practice. *J. Chem. Phys.* 142 (3): 034102.

82 Zgid, D. and Nooijen, M. (2008). Obtaining the two-body density matrix in the density matrix renormalization group method. *J. Chem. Phys.* 128: 144115.

83 Malmqvist, P.-Å., Roos, B.O., and Schimmelpfennig, B. (2002). The restricted active space (RAS) state interaction approach with spin–orbit coupling. *Chem. Phys. Lett.* 357: 230–240.

84 Malmqvist, P.-Å., Pierloot, K., Shahi, A.R.M. et al. (2008). The restricted active space (RAS) followed by second-order perturbation theory method: theory and application to the study of CuO_2 and Cu_2O_2 systems. *J. Chem. Phys.* 128: 204109.

85 Ma, D., Li Manni, G., Olsen, J., and Gagliardi, L. (2016). Second-order perturbation theory for generalized active space self-consistent-field wave functions. *J. Chem. Theory Comput.* 12 (7): 3208–3213.

86 Yaffe, L.G. and Goddard, W.A. (1976). Orbital optimization in electronic wave functions; equations for quadratic and cubic convergence of general multiconfiguration wave function. *Phys. Rev. A* 13: 1682.

87 Roos, B.O., Walch, S.P., Bauschlicher, C.W., and Nelin, C.J. (1983). Theoretical evidence for multiple 3d bonding in the V_2Cr_2 molecules. *Chem. Phys. Lett.* 103: 175.

88 Nakano, H. and Hirao, K. (2000). A quasi-complete active space self-consistent field method. *Chem. Phys. Lett.* 317: 90.

89 Panin, A.I. and Sizova, O.V. (1996). Direct CI method in restricted configuration spaces. *J. Comput. Chem.* 17 (2): 178–184.

90 Panin, A.I. and Simon, K.V. (1996). Configuration interaction spaces with arbitrary restrictions on orbital occupancies. *Int. J. Quantum Chem.* 59: 471.

91 Ivanic, J. (2003). Direct configuration interaction and multiconfigurational self-consistent-field method for multiple active spaces with variable occupations. I. Method. *J. Chem. Phys.* 119: 9364.

92 Fleig, T., Olsen, J., and Marian, C.M. (2001). The generalized active space concept for the relativistic treatment of electron correlation. I. Kramers-restricted two-component configuration interaction. *J. Chem. Phys.* 114: 4775.

93 Liu, Y., Fang, L., Shen, X. et al. (2000). Absorption, resonance Raman and Raman excitation spectra of lanthanum dimers in argon matrices. *Chem. Phys.* 262: 25.

94 Golub, G.H., Luk, F.T., and Overton, M.L. (1981). A block Lanczos method for computing the singular values and corresponding singular vectors of a matrix. *ACM Trans. Math. Softw.* 7 (2): 149–169.

95 Austin, B.M., Zubarev, D.Y., and Lester, W.A. (2012). Quantum Monte Carlo and related approaches. *Chem. Rev.* 112 (1): 263–288. PMID: 22196085.

96 Petruzielo, F.R., Holmes, A.A., Changlani, H.J. et al. (2012). Semistochastic projector Monte Carlo method. *Phys. Rev. Lett.* 109: 230201.

97 Blunt, N.S., Smart, S.D., Kersten, J.A.-F. et al. (2015). Semi-stochastic full configuration interaction quantum Monte Carlo: developments and application. *J. Chem. Phys.* 142: 184107.

98 Blunt, N.S., Smart, S.D., Booth, G.H., and Alavi, A. (2015). An excited-state approach within full configuration interaction quantum Monte Carlo. *J. Chem. Phys.* 143 (13).

99 Blunt, N.S., Booth, G.H., and Alavi, A. (2017). Density matrices in full configuration interaction quantum Monte Carlo: excited states, transition dipole moments, and parallel distribution. *J. Chem. Phys.* 146 (24): 244105.

7

The Density Matrix Renormalization Group for Strong Correlation in Ground and Excited States

Leon Freitag[1] *and Markus Reiher*[2]

[1] *Institute of Theoretical Chemistry, Faculty of Chemistry, University of Vienna, Vienna, Austria*
[2] *ETH Zürich, Laboratorium für Physikalische Chemie, Vladimir-Prelog-Weg 2, 8093 Zürich, Switzerland*

Abstract

The density matrix renormalization group (DMRG), originally introduced by White in 1992 in solid state physics, has since found numerous applications in quantum chemistry. DMRG allows one to approximate the full CI wave function with polynomial scaling, making active spaces with about 100 orbitals accessible. Together with self-consistent field orbital optimization (DMRG-SCF), it allows for much larger active spaces than the complete active space self-consistent field (CASSCF) method. In this chapter, we provide an introduction to the theory behind DMRG, both in the original renormalization group formulation, as well as in the more modern formulation where wave functions and operators are represented as matrix product states and matrix product operators, respectively. We further discuss quantum-information-theoretical orbital entanglement measures that are accessible through DMRG, which pave the way to automated active space selection in multiconfigurational calculations, and factors that control DMRG convergence and accuracy. Finally, we review modern developments in and around DMRG, such as post-DMRG methods for the description of dynamic correlation and environment effects, tensor network states, as well as applications of DMRG and DMRG-SCF in quantum chemistry.

7.1 Introduction

Accurate calculations of ground- and excited-state properties of large molecules and clusters are important applications of modern electronic structure theory. When such systems (and, in particular, their excited states) show strong electron correlation (also referred to as static correlation, left-right correlation, or non-dynamic correlation in the literature), they cannot be adequately described by a single-determinant wave function ansatz, but require a multi-reference method instead.

An important branch of modern multi-reference methods is based on exact or approximate complete active space self-consistent field (CASSCF) wave functions [1, 2] (also called *fully optimized reaction space* (FORS) approximation [3–5]). Construction of a CASSCF wave function requires a definition of an active orbital space, solution of the full configuration interaction (full CI) problem within this active space (which yields an exact solution of the Schrödinger equation within the basis spanned by the active orbitals), and orbital optimization. The grand challenge of CAS-based methods is overcoming the exponential scaling of the CI problem with the size of

Quantum Chemistry and Dynamics of Excited States: Methods and Applications,
First Edition. Edited by Leticia González and Roland Lindh.

the active space. Exact CASSCF, even with the latest hardware and a massively parallel optimized code, are currently limited to about 22 electrons in 22 orbitals [6], which can be a severe limitation for many multi-reference problems.

Therefore, numerous efforts in the development of the multi-reference methods have focused on developing approximations to the full CI problem to allow for calculations with larger active spaces.

The traditional approach to this objective reduces the configuration space (and therefore the dimensions of the CI problem) by restricting the excitations within the active space, leading to methods such as the restricted active space self-consistent field (RASSCF) [7], its extension – the generalized active space SCF (GASSCF) [8], occupation restricted multi-level active space (ORMAS) [9], and related methods, which are described in Chapter 6. A related set of approaches is the Split-CAS [10] and its extension Split-GAS [11] by Gagliardi and co-workers, where the active space is partitioned into a small primary and a larger extended space, the latter one being incorporated with a perturbative treatment. Zimmerman [12] recently presented a novel CI scheme with incremental truncations, approaching full CI accuracy. Similarly, Eriksen et al. [13, 14] recently developed another CI scheme with incremental truncations based on the many-body expansions.

Other approaches sample the configuration space. This includes methods such as full CI Quantum Monte-Carlo (FCIQMC) [15, 16] or heat-bath CI [17, 18], the adaptive sampling CI (ASCI) [19, 20] or the semistochastic FCIQMC, which is related to all the previous methods [21, 22]. Many of these approaches incorporate ideas that are based on the CIPSI method of Malrieu and co-workers [23] and the work of Buenker and Peyerimhoff [24] from the 1970s. Other noteworthy CIPSI-based methods are the difference-dedicated CI (DDCI) method, again by Malrieu and co-workers [25, 26] and the spectroscopically oriented CI (SORCI) by Neese [27].

Another promising research direction aims at finding low-rank approximations to the CI problem by exploiting various (approximate) decompositions of the full CI vector. Olsen et al. [28] have incorporated a decomposition of the CI vector reshaped in a matrix into an approximation of the CI singles and doubles (CISD) method. Koch and Dalgaard [29] considered a decomposition of the CI vector with variationally determined decomposition coefficients. Based on these ideas, a very promising CI approximation named *rank-reduced full CI* has been developed very recently by Martínez and co-workers [30]. Shepard et al. [31–35] have devised the *graphically contracted function CI* (GCF-CI). Although not strictly a decomposition-based method, it employs a reduced number of variational parameters similar to other decomposition-based methods. The variational parameters in GCF-CI arise from the graphical unitary group approach (GUGA) representation [36, 37] of the full CI wave function and span the same space as the full set of the CI coefficients. However, by far the largest class of CI vector decomposition-based approximations are the tensor decomposition methods, in which the CI vector is decomposed into a series of tensors and low-rank approximations are devised for the tensors at hand. The density matrix renormalization group (DMRG) belongs to this group.

Introduced by White [38, 39] in 1992 for one-dimensional computational problems in physics (such as Heisenberg spin chains), DMRG has been applied to quantum chemistry in 1998 with the Pariser-Parr-Pople Hamiltonian to model excited states of conjugated systems [40, 41] and quickly broadened its range of quantum chemical applications with the implementation of the full non-relativistic Coulomb Hamiltonian in the late 1990s and early 2000s by White and co-workers [42, 43], Mitrushenkov et al. [44–46], Chan et al. [47–51], Legeza et al. [52], Reiher et al. [53, 54], Zgid and Nooijen [55–57] and Yanai et al. [58], finding numerous applications in quantum chemistry ever since (see Section 7.5). DMRG is capable of producing variational approximate solutions

to the full CI problem at polynomially scaling costs. It is iterative and introduces approximations, but it has the capability of error control [47, 59–62]. These features have made DMRG (and its related methods) a rising star among the multi-configurational quantum chemical methods, with applications to systems with active orbital spaces even as large as 100 electrons in 100 orbitals [63].

7.2 DMRG Theory

The formulation of the DMRG algorithm may follow either the original *renormalization group* formulation or a formulation based on *matrix product states* and *matrix product operators* (MPS-MPO). In Section 7.2.1 we discuss the concepts of DMRG within the framework of its original formulation. Then, Section 7.2.2 introduces matrix product states and operators, and Section 7.2.3 presents the variational MPS-MPO formulation of DMRG. This chapter is in no way an exhaustive review of DMRG, for which the reader is referred to the reviews of Schollwöck [64, 65] and Hallberg [66, 67] which give an exhaustive introduction into DMRG theory, and to the numerous reviews by Chan et al. [68–72], Marti and Reiher [61, 73] and Wouters and Van Neck [74] on DMRG in quantum chemistry.

A Brief Note on Terminology

In DMRG, orbitals are called *sites* and are arranged on a one-dimensional *lattice*, resembling the one-dimensional nature of the algorithm. This choice of terminology originates from the original formulation of DMRG for a Heisenberg spin chain, where the spins are physically situated on adjacent sites on a one-dimensional lattice. However in quantum chemistry the sites correspond to molecular orbitals which usually do not resemble the one-dimensional structure of the lattice at all. In fact, different orbital orderings on the lattice change the convergence behavior of DMRG, as will be discussed later.

7.2.1 Renormalization Group Formulation

We express a multi-configurational wave function in terms of a superposition of occupation number vectors:

$$|\Psi\rangle = \sum_{\sigma} c_{\sigma} |\sigma\rangle \tag{7.1}$$

The *basis states* $|\sigma\rangle$ (not to be confused with electronic states $|\Psi\rangle$) span all possible occupations of L spatial orbitals and correspond to Slater determinants. These many-particle basis states are constructed as a direct product of basis states of the single orbitals $|\sigma_l\rangle$:

$$|\sigma\rangle = \bigotimes_{l=1}^{L} |\sigma_l\rangle. \tag{7.2}$$

There are four possible single-orbital basis states for each $|\sigma_l\rangle$, corresponding to four possible occupations of an orbital: empty ($|-\rangle$), spin-up ($|\uparrow\rangle$), spin-down ($|\downarrow\rangle$) and doubly-occupied ($|\uparrow\downarrow\rangle$). The full dimension of the vector space spanned by $|\sigma\rangle$ is then 4^L, and therefore grows exponentially with L. For example, for two orbitals we have $4^2 = 16$ basis states, depicted in Table 7.1. Introducing electron number conservation or spin symmetry reduces the number of basis states. For the treatment of symmetries in DMRG we may refer to the reviews of Sharma and Chan [75] and Wouters and Van Neck [74], as well as to our work [76].

Table 7.1 Basis states for a two-orbital lattice consisting of two single-orbital basis states.

	$\|-\rangle$	$\|\uparrow\rangle$	$\|\downarrow\rangle$	$\|\uparrow\downarrow\rangle$
$\|-\rangle$	$\|--\rangle$	$\|-\uparrow\rangle$	$\|-\downarrow\rangle$	$\|-\uparrow\downarrow\rangle$
$\|\uparrow\rangle$	$\|\uparrow-\rangle$	$\|\uparrow\uparrow\rangle$	$\|\uparrow\downarrow\rangle$	$\|\uparrow\uparrow\downarrow\rangle$
$\|\downarrow\rangle$	$\|\downarrow-\rangle$	$\|\downarrow\uparrow\rangle$	$\|\downarrow\downarrow\rangle$	$\|\downarrow\uparrow\downarrow\rangle$
$\|\uparrow\downarrow\rangle$	$\|\uparrow\downarrow-\rangle$	$\|\uparrow\downarrow\uparrow\rangle$	$\|\uparrow\downarrow\downarrow\rangle$	$\|\uparrow\downarrow\uparrow\downarrow\rangle$

Matrix Representation of the Elementary Operators

We consider the non-relativistic quantum chemical electronic Hamiltonian expressed in second quantization,

$$\hat{H} = \sum_{\substack{i,j \\ \tau}} h_{ij} a^\dagger_{i\tau} a_{j\tau} + \frac{1}{2} \sum_{\substack{i,j,k,l \\ \tau,\tau'}} V_{ijkl} a^\dagger_{i\tau} a^\dagger_{k\tau'} a_{l\tau'} a_{j\tau}, \tag{7.3}$$

with the one- and two-electron integrals h_{ij} and V_{ijkl}, respectively, given in the molecular orbital basis. The creation and annihilation operators we denote $a^\dagger_{i\tau}$ and $a_{i\tau}$, respectively, for an orbital i and spin τ. We now construct matrix representations of the elementary creation and annihilation operators operating on a particular basis state. For example, consider a representation of a creation operator $a^\dagger_{i\uparrow}$ operating on a basis state $|\sigma_i\rangle$ belonging to the single-orbital basis: $\{|-\rangle, |\uparrow\rangle, |\downarrow\rangle, |\uparrow\downarrow\rangle\}$, which translates to the following set of matrix operators:

$$a^\dagger_{i\uparrow}|-\rangle = |\uparrow\rangle \quad \Longrightarrow \quad \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \tag{7.4}$$

$$a^\dagger_{i\uparrow}|\uparrow\rangle = 0 \quad \Longrightarrow \quad \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tag{7.5}$$

$$a^\dagger_{i\uparrow}|\downarrow\rangle = |\uparrow\downarrow\rangle \quad \Longrightarrow \quad \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \tag{7.6}$$

$$a^\dagger_{i\uparrow}|\uparrow\downarrow\rangle = 0 \quad \Longrightarrow \quad \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{7.7}$$

A representation of the same operator on a basis state belonging to a basis of L orbitals $|\sigma_{1..L}\rangle = \{|-\rangle, |\uparrow\rangle, |\downarrow\rangle, |\uparrow\downarrow\rangle\}^{\otimes L}$ then reads

$$a^\dagger_{i\uparrow} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}_1 \otimes \cdots \otimes \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}_i \otimes \cdots \otimes \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}_L \tag{7.8}$$

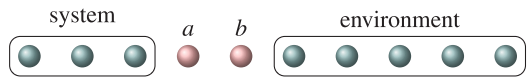


Figure 7.1 An illustration of a lattice partitioning in DMRG into *system*, two *active sites a* and *b* and the *environment* in a microiteration step of a two-site DMRG optimization.

i.e., the operator acts only on site i, leaving other sites untouched. Note that the minus signs in matrices for sites to the left of site i arise from the Jordan–Wigner transformation [77], which is required for fermionic anticommutation. From such matrix representations of elementary operators, a matrix representation of a Hamiltonian acting on any basis state can be constructed by matrix multiplications, weighting with the corresponding integrals and summation of all such terms in the Hamiltonian.

DMRG Iterations

For *two-site* DMRG, the lattice is partitioned into the *system*, two explicitly treated *active sites a* and *b* and the *environment* (see Figure 7.1). We denote the basis spanned by the basis states on the system and environment as $|\text{sys}\rangle$ and $|\text{env}\rangle$ respectively, and the local basis states on a and b as $|a\rangle$ and $|b\rangle$.

We now assume that we can limit the dimension of $|\text{sys}\rangle$ and $|\text{env}\rangle$ by a certain *maximum number of renormalized states m* by some yet undefined procedure (which will be explained later), such that

$$\dim(|\text{sys}\rangle) = \begin{cases} 4^{L_{\text{sys}}}; & 4^{L_{\text{sys}}} < m \\ m; & 4^{L_{\text{sys}}} \geq m \end{cases}$$

$$\dim(|\text{env}\rangle) = \begin{cases} 4^{L_{\text{env}}}; & 4^{L_{\text{env}}} < m \\ m; & 4^{L_{\text{env}}} \geq m \end{cases}$$

with L_{sys} and L_{env} being the number of sites in the system and the environment, respectively.

Next, we may absorb site a into the system and site b into the environment, forming the bases of the enlarged system and the enlarged environment $|i\rangle$ and $|j\rangle$, respectively:

$$|i\rangle = |\text{sys}\rangle \otimes |a\rangle; \qquad |j\rangle = |b\rangle \otimes |\text{env}\rangle$$

with

$$\dim(|i\rangle) = \dim(|\text{sys}\rangle) \cdot \dim(|a\rangle) = 4 \cdot \dim(|\text{sys}\rangle)$$
$$\dim(|j\rangle) = \dim(|\text{env}\rangle) \cdot \dim(|b\rangle) = 4 \cdot \dim(|\text{env}\rangle)$$

so that the dimension of $|i\rangle$ or $|j\rangle$ never exceeds $4m$.

Any state defined on the total orbital space may then be expanded in a tensor product basis composed of enlarged-system and enlarged-environment states:

$$|\Psi\rangle = \sum_{ij} c_{ij} |i\rangle \otimes |j\rangle. \tag{7.9}$$

Consequently, the total dimension of the basis $|i\rangle \otimes |j\rangle$ no longer grows exponentially with the number of sites in the system and environment, but instead does not exceed $m \cdot 4 \cdot 4 \cdot m = 16m^2$.

The reduced-dimensional state is then optimized in the following iterative procedure:

1. Enlarging system and environment by one of the active sites to form basis states $|i\rangle$ and $|j\rangle$, called *blocking*.
2. The *superblock* Hamiltonian (Eq. (7.3)) in this basis is constructed from the elementary operators (cf. Eq. (7.8)) defined in the basis $|i\rangle \otimes |j\rangle$, and then diagonalized. This yields the set

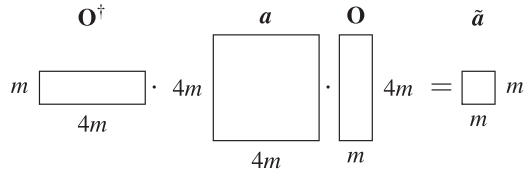


Figure 7.2 An illustration of the transformation of elementary operators $\boldsymbol{a}$ leading to a reduction of dimension. From reduced-dimensional operators ã a reduced-dimensional Hamiltonian is then constructed.

of coefficients c_{ij} in Eq. (7.9). Note that the construction of the matrix representation of the elementary operators is at first performed explicitly as described before, but is adjusted to the reduced-dimensional basis if the decimation step (see step 4) is performed.

3. From the coefficients c_{ij}, a reduced density matrix $\boldsymbol{\rho}^{\mathrm{s/e}}$ of dimension $\dim(|i\rangle) \times \dim(|i\rangle)$, i.e., at most $4m \times 4m$, is formed by tracing out the environment states:

$$\rho^{\mathrm{s/e}}_{ii'} = \sum_{j \in |j\rangle} c_{ij} c_{i'j} \tag{7.10}$$

4. Then, m eigenvectors of $\boldsymbol{\rho}^{\mathrm{s/e}}$ that correspond to the largest eigenvalues are chosen to form a rectangular transformation matrix $\mathbf{O}$ (of dimension $m \times 4m$). By transforming an elementary operator $\boldsymbol{a}$ with this matrix according to $\mathbf{O}^\dagger \boldsymbol{a} \mathbf{O}$, their dimension (and hence, the dimension of the Hamiltonian) is reduced from $4m$ to m. The dimension reduction, called *decimation*, is pictorially illustrated in Figure 7.2. The decimation step ensures that the dimension of the system basis does not grow beyond m. Note that if the dimension of $\boldsymbol{\rho}^{\mathrm{s/e}}$ is less than $m \times m$, no decimation needs to be performed.
   It can be shown [38] that the decimation procedure based on the reduced density matrix yields an optimal approximation $|\tilde{\Psi}\rangle$ to the un-truncated wave function in Eq. (7.9) in a least-squares sense, i.e., it minimizes the error $|\Psi - \tilde{\Psi}|^2$.
5. Once this microiteration step has been accomplished, the active sites are shifted to the right by one site on the lattice (so that the environment shrinks by one site). Steps 1–4 are repeated until the environment consists of only one site. Then, the environment and the system exchange their roles: the new system is growing to the left of the lattice and the active sites are shifted from right to left. A complete set of repeated blocking-optimization-decimation steps along the lattice is called a *sweep*. Figure 7.3 depicts DMRG microiteration steps schematically.

7.2.2 Matrix Product States and Matrix Product Operators

Three years after White [38, 39] published his first papers on DMRG, Östlund and Rommer [79, 80] made a connection of the DMRG optimization algorithm with a wave function form of specific parametrization, the so-called *matrix product states* (MPS). In this and in the following section, we will closely follow Refs. [64, 65] and [81].

A general multi-configurational wave function (cf. Eq. (7.1)) can be brought through a series of singular-value decompositions (SVD) [65] into the following form:

$$|\Psi\rangle = \sum_{\sigma} \mathbf{M}^{\sigma_1} \mathbf{M}^{\sigma_2} \cdots \mathbf{M}^{\sigma_L} |\sigma\rangle \tag{7.11}$$

which after writing the matrix–matrix multiplications explicitly reads

$$|\Psi\rangle = \sum_{\sigma_1 \ldots \sigma_L} \sum_{a_1 \ldots a_{L-1}} M^{\sigma_1}_{1 a_1} M^{\sigma_2}_{a_1 a_2} \cdots M^{\sigma_L}_{a_{L-1} 1} |\sigma_1 \otimes \sigma_2 \otimes \cdots \sigma_L\rangle \tag{7.12}$$

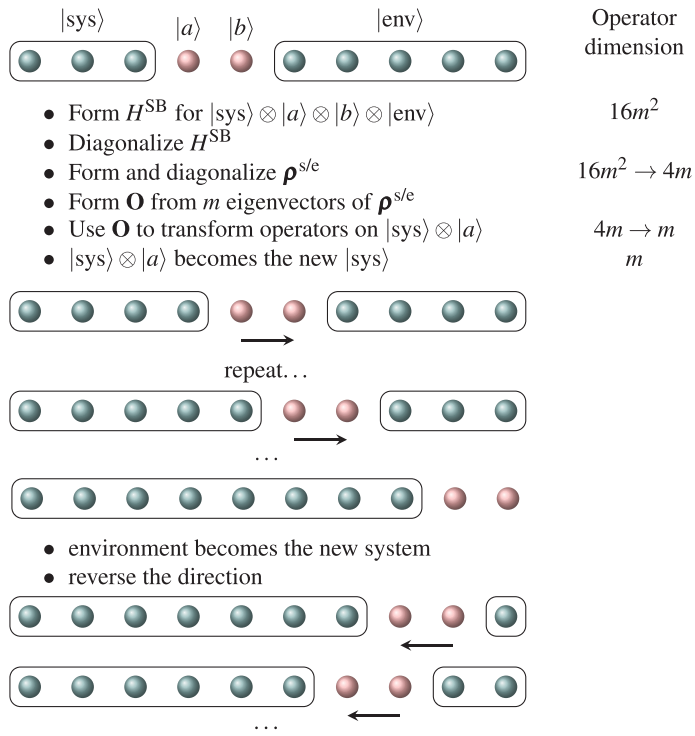


Figure 7.3 Illustration of DMRG microiterations. $|sys\rangle$, $|a\rangle$, $|b\rangle$ and $|env\rangle$ represent the states on the system, sites a and b and the environment, respectively. Adapted from Ref. [78].

where a tensor with elements $M^{\sigma_l}_{a_{l-1}a_l}$ corresponds to each site l. As we are summing over all local basis states $|\sigma_l\rangle$ as defined in Section 7.2.1, so that each σ_l corresponds to a local occupation of a site l, there are four such possible local occupations. Each tensor $\{M^{\sigma_l}_{a_{l-1}a_l}\}$ therefore can be considered as a vector of four matrices with vector elements $\mathbf{M}^{\sigma_l}$. For an exact decomposition of a full CI state, the dimensions of the $\mathbf{M}^{\sigma_l}$ matrices are $1\times 4, 4\times 16, \ldots, 4^{L/2-1}\times 4^{L/2}, 4^{L/2}\times 4^{L/2-1}, \ldots, 4\times 1$. However, with the help of SVD (see below) we can reduce the matrix dimensions to m. Here we will not go into the details of how the MPS structure is derived: for this, the interested reader is referred to Ref. [64].

Singular Value Decomposition

A general $p\times n$ matrix $\mathbf{M}$ may be factorized as

$$\mathbf{M} = \mathbf{USV}^\dagger \tag{7.13}$$

with a $p\times \min(p,n)$ matrix $\mathbf{U}$, diagonal $\min(p,n)\times\min(p,n)$ matrix $\mathbf{S}$ with real non-negative elements, and a $\min(p,n)\times n$ matrix $\mathbf{V}$. Matrices $\mathbf{U}$ and $\mathbf{V}$ have orthogonal columns, i.e., $\mathbf{U}^\dagger\mathbf{U} = \mathbf{V}^\dagger\mathbf{V} = \mathbf{I}$. If we write out the matrix multiplication explicitly, the SVD will read

$$M_{pn} = \sum_{k=\min(p,n)} U_{pk}S_{kk}V^\dagger_{kn}. \tag{7.14}$$

We may build a matrix $\mathbf{M'}$ by setting all but the largest m diagonal elements S_{kk} to zero and using only the first m columns of $\mathbf{U}$ and $\mathbf{V}$, which is referred to as *truncated SVD*. The truncated SVD ensures that $\mathbf{M'}$ provides the best low-rank approximation to $\mathbf{M}$ such that the error

$||\mathbf{M} - \mathbf{M}'||^2 = \sum_{ij}(M_{ij} - M'_{ij})^2$ is minimal. Employing this procedure in the MPS construction ensures that the matrix dimension in the MPS does not exceed m. In Section 7.2.4 we will see that SVD is equivalent to the reduced density matrix diagonalization in the decimation step.

Left- and Right-Normalization

An MPS representation is not unique: if we insert an identity matrix $\mathbf{I} = \mathbf{X}\mathbf{X}^{-1}$ between any two matrices $\mathbf{M}^{\sigma_i}$ and $\mathbf{M}^{\sigma_{i+1}}$ (note that, as above, $\mathbf{M}^{\sigma_i}$ is a vector element of the tensor M at site i, reshaped into a vector of matrices) and thus replace $\mathbf{M}^{\sigma_i}$ and $\mathbf{M}^{\sigma_{i+1}}$ with $\bar{\mathbf{M}}^{\sigma_i} = \mathbf{M}^{\sigma_i}\mathbf{X}$ and $\bar{\mathbf{M}}^{\sigma_{i+1}} = \mathbf{X}^{-1}\mathbf{M}^{\sigma_{i+1}}$, the overall MPS will not change. This property allows us to bring the MPS into a so-called *left-canonical form*

$$|\Psi\rangle = \sum_{\sigma} \mathbf{A}^{\sigma_1}\mathbf{A}^{\sigma_2} \cdots \mathbf{A}^{\sigma_L}|\sigma\rangle \tag{7.15}$$

where the $\mathbf{A}^{\sigma_i}$ matrices are *left-normalized* ($\sum_{\sigma_i=1}^{4} \mathbf{A}^{\sigma_i\dagger}\mathbf{A}^{\sigma_i} = \mathbf{I}$), or into a *right-canonical form*

$$|\Psi\rangle = \sum_{\sigma} \mathbf{B}^{\sigma_1}\mathbf{B}^{\sigma_2} \cdots \mathbf{B}^{\sigma_L}|\sigma\rangle \tag{7.16}$$

where the $\mathbf{B}^{\sigma_l}$ matrices are *right-normalized* ($\sum_{\sigma_i=1}^{4} \mathbf{B}^{\sigma_i}\mathbf{B}^{\sigma_i\dagger} = \mathbf{I}$). A form that is important in the sweep procedure is the *mixed-canonical* form

$$|\Psi\rangle = \sum_{\sigma} \mathbf{A}^{\sigma_1} \cdots \mathbf{A}^{\sigma_{l-1}}\mathbf{M}^{\sigma_l}\mathbf{B}^{\sigma_{l+1}} \cdots \mathbf{B}^{\sigma_L}|\sigma\rangle \tag{7.17}$$

where at an arbitrary site l the matrices to the left are left-normalized and the matrices to the right are right-normalized. These canonical representations will come in handy when calculating overlaps, expectation values and optimizing the MPS.

The left-normalization is usually achieved with an SVD: if one reshapes the tensor M at site l (with the elements $M^{\sigma_l}_{a_{l-1}a_l}$) in Eq. (7.17) (note that the state in Eq. (7.17) must be only partially left-normalized, i.e., tensors at sites $> l$ need not be right-normalized) as one matrix with elements $M_{(\sigma_l a_{l-1}),a_l}$ and performs an SVD, one will obtain

$$M_{(\sigma_l a_{l-1}),a_l} = \sum_{s_l} U_{(\sigma_l a_{l-1}),s_l} S_{s_l,s_l} V^{\dagger}_{s_l,a_l} \tag{7.18}$$

$U_{(\sigma_l a_{l-1}),s_l}$ is then reshaped back into a tensor with elements $A^{\sigma_l}_{a_{l-1}s_l}$ which obeys the left normalization condition, and $\mathbf{SV}^{\dagger}$ is pre-multiplied into the matrices at site $l+1$.

Right-normalized matrices are obtained in a similar manner by starting from a partially right-normalized state: this time, we group the σ_l index with a_l, resulting in an SVD

$$M_{a_{l-1},(\sigma_l a_l)} = \sum_{s_l} U_{s_l} S_{s_l,s_l} V^{\dagger}_{s_l,(\sigma_l a_l)} \tag{7.19}$$

Now $V^{\dagger}_{s_l,(\sigma_l a_l)}$ is reshaped back into a tensor $B^{\sigma_l}_{a_{l-1}s_l}$ which obeys the right normalization condition and $\mathbf{US}$ is post-multiplied into the matrices at site $l-1$.

In principle, one may use the QR decomposition to perform the normalization instead of SVD. However, with the QR decomposition it is not possible to perform truncation, which, although not required in the normalization step, may be required in the sweep algorithm: more details will follow at the end of Section 7.2.3.

Overlaps

An overlap of two MPSs is straightforward to calculate. Given the state

$$|\bar{\Psi}\rangle = \sum_{\sigma} \sum_{a_1 \ldots a_{L-1}} \bar{M}^{\sigma_1}_{1a_1} \bar{M}^{\sigma_2}_{a_1 a_2} \cdots \bar{M}^{\sigma_L}_{a_{L-1}1} |\sigma\rangle, \tag{7.20}$$

the overlap can be calculated as

$$\langle \Psi | \bar{\Psi} \rangle = \sum_{\substack{a'_1,\ldots,a'_{L-1} \\ a_1,\ldots,a_{L-1} \\ \sigma}} (M^{\sigma_1}_{1a_1} \cdots M^{\sigma_L}_{a_{L-1}1})^* (\bar{M}^{\sigma_1}_{1a'_1} \cdots \bar{M}^{\sigma_L}_{a'_{L-1}1}) \tag{7.21}$$

or, regrouped for a more efficient computation

$$\langle \Psi | \bar{\Psi} \rangle = \sum_{\substack{a_{L-1} a'_{L-1} \\ \sigma_L}} M^{\sigma_L \dagger}_{1a_{L-1}} \left(\ldots \sum_{\substack{a_1 a'_1 \\ \sigma_2 \sigma'_2}} M^{\sigma_2 \dagger}_{a_2 a_1} \left(\sum_{\sigma_1} M^{\sigma_1 \dagger}_{a_1 1} \bar{M}^{\sigma_1}_{1a'_1} \right) \bar{M}^{\sigma_2}_{a'_1 a'_2} \cdots \right) \bar{M}^{\sigma_L}_{a'_{L-1}1}. \tag{7.22}$$

Matrix Product Operators

Analogously to an MPS, an operator $\hat{\mathcal{W}}$ of general form

$$\hat{\mathcal{W}} = \sum_{\sigma,\sigma'} c_{\sigma,\sigma'} |\boldsymbol{\sigma}\rangle\langle\boldsymbol{\sigma'}| \tag{7.23}$$

may be brought to the *matrix product operator* (MPO) form

$$\hat{\mathcal{W}} = \sum_{\sigma,\sigma'} \sum_{b_1 \ldots b_{L-1}} W^{\sigma_1 \sigma'_1}_{1b_1} W^{\sigma_2 \sigma'_2}_{b_1 b_2} \cdots W^{\sigma_L \sigma'_L}_{b_{L-1}1} |\boldsymbol{\sigma}\rangle\langle\boldsymbol{\sigma'}| \tag{7.24}$$

To understand, how an MPO is constructed, let us simplify Eq. (7.24) by introducing

$$\hat{W}^{[l]}_{b_{l-1}b_l} = \sum_{\sigma_l,\sigma_{l'}} W^{\sigma_l \sigma_{l'}}_{b_{l-1}b_l} |\boldsymbol{\sigma_l}\rangle\langle\boldsymbol{\sigma'_l}| \tag{7.25}$$

so that Eq. (7.24) becomes

$$\hat{\mathcal{W}} = \sum_{b_1 \ldots b_{L-1}} \hat{W}^{[1]}_{1b_1} \hat{W}^{[2]}_{b_1 b_2} \cdots \hat{W}^{[L]}_{b_{L-1}1} \tag{7.26}$$

$$= \hat{\mathbf{W}}^{[1]} \hat{\mathbf{W}}^{[2]} \cdots \hat{\mathbf{W}}^{[L]} \tag{7.27}$$

Matrices $\hat{\mathbf{W}}^{[l]}$ are operator-valued, i.e., each matrix element $\hat{W}^{[l]}_{b_{l-1}b_l}$ is an operator. Moreover, matrices $\hat{\mathbf{W}}^{[l]}$ collect all operators which act on site l in matrix form: to see this, we may write, for example, the creation operator $a^\dagger_{l\uparrow}$ in the form analogous to Eq. (7.23), using the basis states at site l:

$$a^\dagger_{l\uparrow} = |\uparrow\downarrow\rangle \langle\downarrow| + |\uparrow\rangle \langle -| \tag{7.28}$$

The equivalence of this form of the operator to its matrix form may be easily verified by inserting the above equation into Eqs. (7.4) to (7.7) and obtaining the matrix form. Hence, the $a^\dagger_{l\uparrow}$ operator may be a matrix element $\hat{W}^{[l]}_{b_{l-1}b_l}$ of the matrix $\hat{\mathbf{W}}^{[l]}$ *for a particular index pair* b_{l-1}, b_l. But how do we determine the indices b_{l-1}, b_l and the dimensions of the matrix $\hat{\mathbf{W}}^{[l]}$? It turns out that determining these indices is not trivial and there is not a unique way to determine them. Construction of MPOs, including an optimal MPO construction algorithm is described thoroughly in Ref. [81].

Expectation Values

The expectation value of any operator reads in the MPS-MPO form as follows:

$$\langle\Psi|\hat{\mathcal{W}}|\bar{\Psi}\rangle = \sum_{\substack{a'_1,\ldots,a'_{L-1}\\ a_1,\ldots,a_{L-1}\\ \sigma\sigma'}} (M^{\sigma_1}_{1a_1}\cdots M^{\sigma_L}_{a_{L-1}1})^* \sum_{b_1,\ldots,b_{L-1}} W^{\sigma_1\sigma'_1}_{1b_1}\cdots W^{\sigma_L\sigma'_L}_{b_{L-1}1}(\bar{M}^{\sigma'_1}_{1a'_1}\cdots \bar{M}^{\sigma'_L}_{a'_{L-1}1}) \tag{7.29}$$

The sum may be regrouped to (analogously to Eq. (7.22)),

$$\begin{aligned}\langle\Psi|\hat{\mathcal{W}}|\bar{\Psi}\rangle = \sum_{\substack{a_{L-1}a'_{L-1}b_{L-1}\\ \sigma_L\sigma'_L}} M^{\sigma_L *}_{1a_{L-1}} W^{\sigma_L\sigma'_L}_{b_{L-1}1} \Bigg(\ldots \sum_{\substack{a_1a'_1b_1\\ \sigma_2\sigma'_2}} M^{\sigma_2 *}_{a_2a_1} W^{\sigma_2\sigma'_2}_{b_1b_2} \times \\ \times \Bigg(\sum_{\sigma_1\sigma'_1} M^{\sigma_1 *}_{a_1 1} W^{\sigma_1\sigma'_1}_{1b_1} \bar{M}^{\sigma'_1}_{1a'_1} \Bigg) \bar{M}^{\sigma'_2}_{a'_1a'_2} \cdots \Bigg) \bar{M}^{\sigma'_L}_{a'_{L-1}1},\end{aligned} \tag{7.30}$$

where we begin contracting the tensors from the first site and continue to the right. This can be turned into a recursive equation with

$$\mathbb{L}^{b_0}_{a_0a'_0} = 1, \tag{7.31}$$

$$\mathbb{L}^{b_1}_{a_1a'_1} = \sum_{\sigma_1\sigma'_1} M^{\sigma_1 *}_{a_1 1} W^{\sigma_1\sigma'_1}_{1b_1} \bar{M}^{\sigma'_1}_{1a'_1}, \tag{7.32}$$

$$\mathbb{L}^{b_l}_{a_la'_l} = \sum_{\substack{a_{l-1}a'_{l-1}b_{l-1}\\ \sigma_l\sigma'_l}} M^{\sigma_l *}_{a_la_{l-1}} W^{\sigma_l\sigma'_l}_{b_{l-1}b_l} \mathbb{L}^{b_{l-1}}_{a_{l-1}a'_{l-1}} \bar{M}^{\sigma'_l}_{a'_la'_{l-1}}, \tag{7.33}$$

$$\mathbb{L}^{b_L}_{a_La'_L} = \langle\Psi|\hat{\mathcal{W}}|\bar{\Psi}\rangle. \tag{7.34}$$

We refer to $\mathbb{L}$ as the *left boundary*. Similarly, a right boundary $\mathbb{R}$ may be defined as

$$\mathbb{R}^{b_{l-1}}_{a'_{l-1}a_{l-1}} = \sum_{\substack{a_la'_lb_{l-1}\\ \sigma_l\sigma'_l}} \bar{M}^{\sigma_l}_{a'_{l-1}a'_l} W^{\sigma_l\sigma'_l}_{b_{l-1}b_l} \mathbb{R}^{b_l}_{a'_la_l} M^{\sigma_l *}_{a_la_{l-1}} \tag{7.35}$$

by starting the tensor contraction in Eq. (7.30) at the last site instead of the first and proceeding to the left. The expectation value $\langle\Psi|\hat{\mathcal{W}}|\bar{\Psi}\rangle$ may then be calculated at any site l from left and right boundaries as

$$\langle\Psi|\hat{\mathcal{W}}|\bar{\Psi}\rangle = \sum_{a_la'_lb_l} \mathbb{L}^{b_l}_{a_la'_l} \mathbb{R}^{b_l}_{a'_la_l}. \tag{7.36}$$

The notion of the left and right boundaries, especially their recursive definition, is tightly connected with a sweep in the MPS-MPO formulation of DMRG, described in the next Section. Left and right boundaries are usually constructed from the left- and right-normalized matrices, respectively, which is a requirement for the derivation of the eigenvalue equations used in the sweep procedure.

7.2.3 MPS-MPO Formulation of DMRG

With the help of MPS and MPO, it is possible to express the DMRG algorithm as a variational problem. The ground state search may be formulated as a sequential variational optimization of the

expectation value of the superblock Hamiltonian $\langle\Psi|\hat{H}|\Psi\rangle$ with respect to elements of the matrices $\mathbf{M}^{\sigma_l}$ at one site l at a time, under the constraint that Ψ is normalized, i.e., as an optimization of the following Lagrangian functional with a constraint [82]:

$$\mathcal{L} = \langle\Psi|\hat{H}|\Psi\rangle - \lambda(\langle\Psi|\Psi\rangle - 1). \tag{7.37}$$

To arrive at the eigenvalue equation, we first insert Eq. (7.34) into Eq. (7.36):

$$\langle\Psi|\hat{\mathcal{H}}|\Psi\rangle = \sum_{a_l a'_l b_l} \sum_{\substack{a_{l-1} a'_{l-1} b_{l-1} \\ \sigma_l \sigma'_l}} M^{\sigma_l}_{a_l a_{l-1}} W^{\sigma_l \sigma'_l}_{b_{l-1} b_l} \mathbb{L}^{b_{l-1}}_{a_{l-1} a'_{l-1}} M^{\sigma'_l}_{a'_l a'_{l-1}} \mathbb{R}^{b_l}_{a'_l a_l}. \tag{7.38}$$

Taking the first derivative of Eq. (7.37) with respect to $M^{\sigma_l}_{a_{l-1} a_l}$ (employing Eqns. (7.22) and (7.38)), assuming $\mathbb{L}$ and $\mathbb{R}$ are formed from the left- and right-normalized matrices, respectively, and setting it to zero, we arrive at the eigenvalue equation

$$\sum_{\substack{a_{l-1} a'_{l-1} b_{l-1} b_l \\ \sigma'_l}} W^{\sigma_l \sigma'_l}_{b_{l-1} b_l} \mathbb{L}^{b_{l-1}}_{a_{l-1} a'_{l-1}} M^{\sigma'_l}_{a'_l a'_{l-1}} \mathbb{R}^{b_l}_{a'_l a_l} = \lambda M^{\sigma_l}_{a_{l-1} a_l}. \tag{7.39}$$

We may reshape it into a canonical eigenvalue problem $Hv = \lambda v$ by setting

$$H_{(\sigma_l a_{l-1} a_l)(\sigma'_l a'_{l-1} a'_l)} = \sum_{b_{l-1} b_l} W^{\sigma_l \sigma'_l}_{b_{l-1} b_l} \mathbb{L}^{b_{l-1}}_{a_{l-1} a'_{l-1}} \mathbb{R}^{b_l}_{a'_l a_l} \tag{7.40}$$

and

$$v_{(\sigma_l a_{l-1} a_l)} = M^{\sigma_l}_{a_{l-1} a_l}. \tag{7.41}$$

Sweeping in the MPS-MPO Formulation

The sweep procedure in the variational DMRG formulation then proceeds as follows. Starting from a guess MPS in the right-canonical form, we calculate all expressions for the right boundaries $\mathbb{R}$. Then, from site $l = 1$ to $L - 1$, we need to

- solve the eigenvalue equation (7.39) for $\mathbf{M}^{\sigma_l}$
- left-normalize $\mathbf{M}^{\sigma_l}$ – obtain $\mathbf{A}^{\sigma_l}$ from $\mathbf{U}$ and multiply $\mathbf{SV}^\dagger$ into $\mathbf{M}^{\sigma_{l+1}}$.
- calculate $\mathbb{L}^{b_l}_{a'_l a_l}$ and move one site to the right on the lattice.

At site $l = L$, the procedure is reversed, so that for sites $l = L$ to 2 we proceed as follows:

- solve the eigenvalue equation (7.39) for $\mathbf{M}^{\sigma_l}$
- right-normalize $\mathbf{M}^{\sigma_l}$ – obtain $\mathbf{B}^{\sigma_l}$ from $\mathbf{V}^\dagger$ and multiply $\mathbf{US}$ into $\mathbf{M}^{\sigma_{l-1}}$.
- calculate $\mathbb{R}^{b_l}_{a'_l a_l}$ and move one site to the left on the lattice.

The sweep procedure is schematically illustrated in Figure 7.4 for one-site DMRG (see below). Due to the normalization procedure, the wave function is always in the mixed-canonical form.

One-Site and Two-Site DMRG

The above MPS-MPO formulation has been described for the optimization of a single site only, whereas the DMRG algorithm in the renormalization group formulation, as described in Section 7.2.1, features two sites. It is, however, possible, to modify the algorithm to optimize two sites simultaneously by introducing two modifications to the single-site algorithm. The first modification is introduced in the eigenvalue equation. We introduce the two-site MPS tensor

$$P^{\sigma_l \sigma_{l+1}}_{a_{l-1} a_{l+1}} = \sum_{a_l} M^{\sigma_l}_{a_{l-1} a_l} M^{\sigma_{l+1}}_{a_l a_{l+1}} \tag{7.42}$$

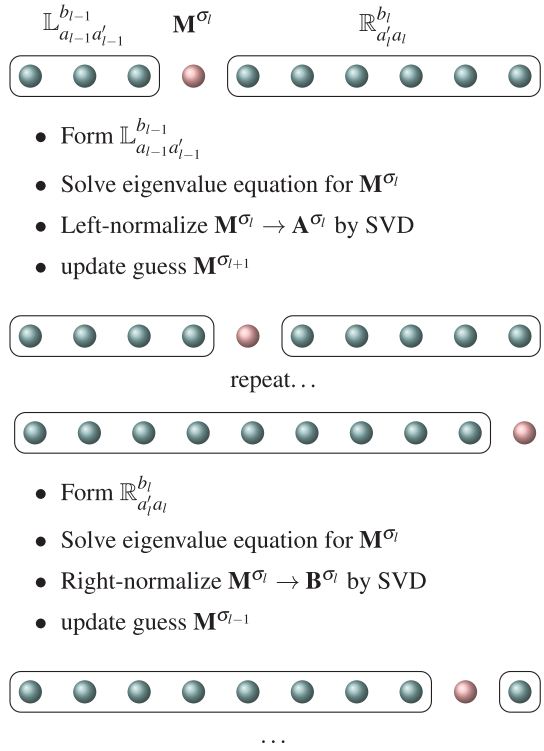


Figure 7.4 Illustration of sweeps for one-site DMRG in the MPS-MPO formulation.

and the two-site MPO tensor

$$\hat{V}^{\sigma_l\sigma_{l+1}\sigma'_l\sigma'_{l+1}}_{b_{l-1}b_{l+1}} = \sum_{b_l} \hat{W}^{\sigma_l\sigma'_l}_{b_{l-1}b_l} \hat{W}^{\sigma_{l+1}\sigma'_{l+1}}_{b_l b_{l+1}} \tag{7.43}$$

and substitute the corresponding tensors in Eq. (7.39). If we consider the local spaces $\sigma_l\sigma_{l+1}$ as a single 16-dimensional local space $\tau_{l,l+1}$, the new eigenvalue equation will become

$$\sum_{\substack{a'_{l-1}a'_{l+1}b_{l-1}b_{l+1}\\ \tau'_{l,l+1}}} V^{\tau_{l,l+1}\tau'_{l,l+1}}_{b_{l-1}b_{l+1}} \mathbb{L}^{b_{l-1}}_{a_{l-1}a'_{l-1}} P^{\tau'_{l,l+1}}_{a'_{l-1}a'_{l+1}} \mathbb{R}^{b_{l+1}}_{a'_{l+1}a_{l+1}} = \lambda P^{\tau_{l,l+1}}_{a_{l-1}a_{l+1}}. \tag{7.44}$$

The second modification to the single-site DMRG algorithm affects the normalization step. Instead of left- or right-normalization, after the solution of the eigenvalue equation, the two-site tensor is again split into $M^{\sigma_l}_{a_{l-1}a_l}$ and $M^{\sigma_{l+1}}_{a_l a_{l+1}}$ with the help of the truncated SVD. We first reshape $P^{\sigma_l\sigma_{l+1}}_{a_{l-1}a_{l+1}}$ into a matrix $P_{(\sigma_l a_{l-1})(\sigma_{l+1}a_{l+1})}$ and then perform the SVD

$$P_{(\sigma_l a_{l-1})(\sigma_{l+1}a_{l+1})} = \sum_{s_l} U_{(\sigma_l a_{l-1})s_l} S_{s_l s_l} V^\dagger_{s_l(\sigma_{l+1}a_{l+1})} \tag{7.45}$$

To construct $M^{\sigma_l}_{a_{l-1}a_l}$ and $M^{\sigma_{l+1}}_{a_l a_{l+1}}$, we retain only m singular values in $\mathbf{S}$, yielding $S_{a_l a_l}$. Now, the M matrices are obtained by reshaping $\mathbf{U}$ and $\mathbf{SV}^\dagger$:

$$M^{\sigma_l}_{a_{l-1}a_l} = U_{(\sigma_l a_{l-1})s_l} \tag{7.46}$$

$$M^{\sigma_{l+1}}_{a_l a_{l+1}} = \frac{1}{\sum_{\tilde{a}_l=1}^{m} S_{\tilde{a}_l\tilde{a}_l}} S_{a_l a_l} V^\dagger_{a_l(\sigma_{l+1}a_{l+1})} \tag{7.47}$$

where in Eq. (7.47) the matrix elements have been scaled with the sum of the retained singular values $\sum_{\tilde{a}_l} S_{\tilde{a}_l \tilde{a}_l}$ for normalization. From the sum of the retained singular values we may derive the truncation error ε [60]

$$\varepsilon = 1 - \sum_{\tilde{a}_l=1}^{m} S_{\tilde{a}_l \tilde{a}_l} \tag{7.48}$$

which allows us to estimate the accuracy of the calculation.

Note how we perform truncation here in the two-site DMRG algorithm only: in the single-site algorithm, there is no truncation, but only a left- or right-normalization step at this point. This turns out to be a large disadvantage of the single-site algorithm. In the two-site algorithm, at each truncation step, m and therefore the dimension of the a_l and a_{l+1} indices may be updated: the basis in which the variational search during the DMRG optimization is performed, therefore, is also optimized. This guarantees variational flexibility and avoids local energy minima. In the single-site algorithm, m is determined *prior to the optimization* during the construction of the MPS guess and is kept fixed during the whole optimization. This leads to a slower convergence and a large tendency to get stuck in local minima [57]. It is, however, possible to introduce truncation at the normalization step also in single-site DMRG by constructing a reduced density matrix as in the renormalization group formulation (cf. Eq. (7.53)), adding noise [47] or a perturbative correction [83] to it and performing the truncation based on the perturbed reduced density matrix. The majority of quantum chemical calculations, however, employ the two-site DMRG algorithm due to the aforementioned better convergence properties. The two-site algorithm is also the closest to the original renormalization group formulation.

7.2.4 Connection between the Renormalization Group and the MPS-MPO Formulation of DMRG

Consider an MPS in a mixed-canonical representation with a two-site tensor at sites l and $l+1$:

$$|\Psi\rangle = \sum_{\sigma, a_1, \ldots, a_{L-1}} A^{\sigma_1}_{1 a_1} \cdots A^{\sigma_{l-1}}_{a_{l-2} a_{l-1}} P^{\sigma_l \sigma_{l+1}}_{a_{l-1} a_{l+1}} B^{\sigma_{l+2}}_{a_{l+1} a_{l+2}} \cdots B^{\sigma_L}_{a_{L-1} 1} |\sigma\rangle. \tag{7.49}$$

We define MPSs that represent two states on the sublattices $\sigma_1 \ldots \sigma_{l-1}$ and $\sigma_{l+2} \ldots \sigma_L$

$$|a_{l-1}\rangle = \sum_{\sigma_1 \ldots \sigma_{l-1}, a_1 \ldots a_{l-2}} A^{\sigma_1}_{1 a_1} \cdots A^{\sigma_{l-1}}_{a_{l-2} a_{l-1}} |\sigma_1 \ldots \sigma_{l-1}\rangle \tag{7.50}$$

$$|b_{l+1}\rangle = \sum_{\sigma_{l+2} \ldots \sigma_L, a_{l+2} \ldots a_L} B^{\sigma_{l+2}}_{a_{l+1} a_{l+2}} \cdots B^{\sigma_L}_{a_{L-1} 1} |\sigma_{l+2} \ldots \sigma_L\rangle \tag{7.51}$$

and insert Eqs. (7.50) and (7.51) into Eq. (7.49). With the relation $|\sigma\rangle = |\sigma_1 \ldots \sigma_{l-1}\rangle \otimes |\sigma_l\rangle \otimes |\sigma_{l+1}\rangle \otimes |\sigma_{l+2} \ldots \sigma_L\rangle$, we arrive at

$$|\Psi\rangle = \sum_{\sigma_l, \sigma_{l+1}, a_{l-1}, a_{l+1}} P^{\sigma_l \sigma_{l+1}}_{a_{l-1} a_{l+1}} |a_{l-1}\rangle \otimes |\sigma_l\rangle \otimes |\sigma_{l+1}\rangle \otimes |b_{l+1}\rangle \tag{7.52}$$

We immediately recognize the equivalence of Eq. (7.52) and Eq. (7.9) and the lattice partitioning from Figure 7.1: the sublattices $\sigma_1 \ldots \sigma_{l-1}$ and $\sigma_{l+2} \ldots \sigma_L$ and sites l and $l+1$ correspond to the system, environment and the active sites a and b, respectively. The mixed-canonical representation ensures the orthogonality of the bases $|a_{l-1}\rangle$ and $|b_{l+1}\rangle$.

The solution of the eigenvalue equation (7.44) is equivalent to diagonalizing the Hamiltonian and obtaining coefficients c_{ij} in Eq. (7.9): if we reshape the two-site tensor $P^{\sigma_l \sigma_{l+1}}_{a_{l-1} a_{l+1}}$ into a matrix $\mathbf{P}$

with matrix elements $P_{(\sigma_l a_{l-1})(\sigma_{l+1} a_{l+1})}$, we see that the first combined index $(\sigma_l a_{l-1})$ corresponds to i in c_{ij} and the second combined index $(\sigma_{l+1} a_{l+1})$ corresponds to j. The reduced density matrix $\boldsymbol{\rho}^{\mathrm{s/e}}$ is simply given as the product $\mathbf{P}\mathbf{P}^\dagger$. By definition of the matrix multiplication we obtain

$$(\mathbf{P}\mathbf{P}^\dagger)_{ii'} = \sum_j P_{ij} P^\dagger_{ji'} \tag{7.53}$$

which is recognized as equivalent to Eq. (7.10). Inserting the SVD of $\mathbf{P}$ in this equation

$$\mathbf{P}\mathbf{P}^\dagger = \mathbf{U}\mathbf{S}\mathbf{V}^\dagger(\mathbf{U}\mathbf{S}\mathbf{V}^\dagger)^\dagger = \mathbf{U}\mathbf{S}\mathbf{V}^\dagger\mathbf{V}\mathbf{S}^\dagger\mathbf{U}^\dagger = \mathbf{U}\mathbf{S}\mathbf{S}^\dagger\mathbf{U}^\dagger = \mathbf{U}\boldsymbol{\lambda}\mathbf{U}^\dagger$$

we arrive at the eigendecomposition of the reduced density matrix $\mathbf{P}\mathbf{P}^\dagger$. Hence, the matrix $\mathbf{U}$ which contains eigenvectors of the reduced density matrix (and which is used in the decimation step in the renormalization group formulation of DMRG) can be equally obtained with the SVD of $\mathbf{P}$.

The left- and right-boundary propagation in the MPS-MPO formulation correspond to basis transformations of the operators and to the shift of the system and the environment blocks in the renormalization group formulation. However, these procedures are not strictly equivalent: in the renormalization group formulation, the operator representation (and, in particular, the Hamiltonian) are always dependent on the basis of the system and the environment block. In the MPS-MPO formulation the operators are constructed independently of the wave function optimization, i.e., the b_i indices in the MPO representation (cf. e.g., Eq. (7.24)) are determined during the MPO construction and remain fixed throughout the whole optimization process. This is the main advantage of the MPS-MPO formulation: the MPO construction is completely independent of the sweep procedure, which allows for easier implementation of various operators, as well as the calculation of expectation values of arbitrary operators independently of the wave function optimization.

7.2.5 Developments to Enhance DMRG Convergence and Performance

To accelerate DMRG calculations, and to gain access to even larger active orbital spaces, several approximations to DMRG have been devised. Quadratic scaling DMRG [63, 84] exploits the locality of the two-electron integrals in a localized basis and employs pre-screening, which removes renormalized operators based on integrals that are on sites too far away from each other on the lattice and vanish. The method works particularly well for elongated systems, where the integrals decay quickly in one dimension.

Ren et al. [85] devised the *inner space perturbation theory*, which replaces the exact optimization of $\mathbf{M}^{\sigma_l}$ by an approximate solution with perturbative correction. A conceptually similar approach is followed in *perturbative DMRG* (p-DMRG) of Guo et al. [86] Here, a perturbative correction is applied on top of a converged DMRG wave function obtained for a small m.

All of the three algorithms improve the convergence compared to standard DMRG, especially for small m values, but at a price of not being fully variational anymore.

7.3 DMRG and Orbital Entanglement

The DMRG algorithm provides means for characterizing the wave function in terms of measures from quantum information theory [87]. Consider a DMRG bipartition of the system (Eq. (7.9))

where the system $|i\rangle$ consists only of one orbital and the environment $|j\rangle$ incorporates all other orbitals. The reduced density matrix (Eq. (7.10)) is then referred to as a *one-orbital* RDM (1o-RDM), and it is possible to define a von-Neumann-type entropy, namely the *single-orbital entropy* $s(1)_i$ based on its eigenvalues $\omega_{\alpha,i}$ [88–91]:

$$s(1)_i = -\sum_{\alpha=1}^{4} \omega_{\alpha,i} \ln \omega_{\alpha,i} \tag{7.54}$$

Since $|i\rangle$ refers to only one spatial orbital, the dimension of the 1o-RDM is 4×4, corresponding to the four single-orbital basis states $\{|-\rangle, |\uparrow\rangle, |\downarrow\rangle, |\uparrow\downarrow\rangle\}$, and α therefore runs over the four eigenvalues of the 1o-RDM.

The single-orbital entropy quantifies the deviation of the state $|i\rangle$ from a pure state (corresponding to one of the four local basis states), and, at the same time, the entanglement of the orbital i with all other orbitals, i.e., its contribution to the multi-configurational character of the state. If the orbital i shows the same occupation in all Slater determinants of the total wave function, the state $|i\rangle$ will be pure and correspond to that occupation, yielding a zero $s(1)_i$. If the occupation of orbital i differs for various Slater determinants, the state $|i\rangle$ will have contributions from different basis states and will, therefore, deviate from a pure state, leading to a non-zero $s(1)_i$.

Analogously, we may introduce a two-orbital entropy $s(2)_{ij}$ for a pair of orbitals i and j by extending the system in Eq. (7.9) to two orbitals and constructing the *two-orbital* reduced density matrix (2o-RDM):

$$s(2)_{ij} = -\sum_{\alpha=1}^{16} \omega_{\alpha,ij} \ln \omega_{\alpha,ij} \tag{7.55}$$

where α runs over the 16 eigenvalues of the 2o-RDM, corresponding to 16 possible basis states for two orbitals. Subtracting the two-orbital entropy from the two single-orbital entropies, we obtain the mutual information I_{ij}

$$I_{ij} = \frac{1}{2}(s(1)_i + s(1)_j - s(2)_{ij})(1 - \delta_{ij}) \tag{7.56}$$

where we follow the definition of I_{ij} from Rissler et al. [89], noting that a definition with an inverted sign can also be found in the literature (see, e.g., Refs. [92, 93]).

Originally introduced for improving the convergence of DMRG calculations, single-orbital entropy and mutual information have since found numerous applications in theoretical physics and chemistry [94]. Examples are the characterization of static and dynamic correlation in molecules with a multi-configurational character [90, 95] and the automated active space selection for multi-configurational calculations [96–98]. We have successfully applied the automated active space selection to several multi-configurational studies: characterization of novel dinuclear Ir(IV,V) complexes [99], elucidation of a N–I bond activation, a key step in the catalytic cycle of a photoactivated iodine-mediated C–H amination [100], multi-configurational perturbation theory calculations on dissociation reactions of transition metal complexes [101], and for the exploration of reaction coordinates and excited states [102]. In addition, we have developed a multi-reference diagnostic based on the orbital entanglement measures [103].

Note that it is also possible to define the 1o-RDM and 2o-RDM in terms of the traditional n-particle reduced density matrices [91–93] and may, in principle therefore, be obtained with any quantum chemical method.

7.4 DMRG in Practice

7.4.1 Calculating Excited States with DMRG

The original ansatz for calculating excited states in the traditional formulation involves the calculation of several eigenstates of the superblock Hamiltonian and employing an average reduced density matrix for all states (Eq. (7.10)) for the decimation step [66, 74, 104]. This algorithm is referred to as *state-averaged DMRG* (not to be confused with the state-averaged DMRG-SCF, where the average orbitals are optimized, just as in state-averaged CASSCF). Alternatively, in the MPS-MPO formalism one may employ a state-specific algorithm, which optimizes each state individually, but orthogonalizes the solution of the eigenvalue problem against lower excited states [81, 105]. Further algorithms include the harmonic Davidson algorithm [84], which is able to target specific excited states or regions in the spectrum, or calculate excited states with the linear response formalism [104].

7.4.2 Factors Affecting the DMRG Convergence and Accuracy

Initial Guess

The DMRG algorithm may converge with a random initial guess, but its convergence may be far from optimal. By applying several techniques, one may accelerate the convergence and decrease the required number of renormalized states m [88, 106] to reach convergence. As mentioned in Section 7.2.3, one, for example, may add noise [47] or a perturbative correction [83] to the reduced density matrix (Eq. (7.10) or (7.53)) before renormalization to prevent the algorithm from getting stuck in a local minimum and accelerate convergence. Alternatively, a guess may be obtained from a chemically sound wave function, e.g., consisting of one or several Slater determinants [107]. Legeza and Sólyom [88] have developed CI-DEAS (CI-guided dynamically extended active space) – a procedure which constructs a guess with the most important Slater determinants based on entanglement information from a preliminary DMRG calculation.

Orbital Choice and Ordering

Due to the inherent structure of an MPS and the sweep procedure that optimizes the sites sequentially, the convergence of DMRG is greatly affected by the orbital ordering. Legeza and Sólyom [88] have suggested using the ordering based on quantum information measures and to place the strongly correlated orbitals next to each other and towards the middle of the lattice. One may calculate a weighted graph from the sites and their mutual information and construct a so-called Fiedler vector [108] of this graph, which can then be exploited for the orbital ordering. A Fiedler ordering may, however, lead to a poorer convergence than even a random ordering if the orbitals are not ordered in groups of the same irreducible representations [108]. Alternatively, instead of exploiting the mutual information, one may base the Fiedler ordering on the absolute values of the exchange integral matrix elements between orbital pairs [71]. As an alternative for the Fiedler ordering, genetic-algorithm-based ordering has been proposed [53, 71].

The DMRG convergence (and energy), of course, also depends on the kind of orbitals employed in the calculation. For this reason, just as CI in a truncated orbital space, DMRG may be combined with the orbital optimization procedure, resulting in the DMRG self-consistent field method (DMRG-SCF) [56, 109–111]. However, in general, a DMRG wave function is not invariant under orbital rotations (unless a large m is used), and hence the energy also depends on which types of orbitals are employed. In addition to canonical or natural orbitals, localized or split-localized orbitals are often used [46, 47, 51, 63, 71].

Number of Renormalized States

The number of renormalized states m is the principal parameter in DMRG that controls the accuracy and the computational resources required. Hence, choosing the right m is always a tradeoff between speed and accuracy. Typically, a value of m between 100 and 10000 is used for the electronic structure calculations, depending on active space size, initial guess, and the orbital ordering.

Unfortunately, the entanglement generally increases with the size of the orbital space and therefore m must be also increased to retain the same accuracy.

The truncation error ϵ (Eq. (7.48)) may be used to estimate the accuracy of a DMRG calculation. As the truncation error decreases with the increasing m, it is possible to perform DMRG calculations with several different m values and obtain an extrapolated energy E_{ext} for a truncation error limit of zero, which corresponds to $m \to \infty$ with a function fit. The following fit function has been suggested by Legeza et al. [60]:

$$\ln \frac{E_{\text{DMRG}} - E_{\text{ext}}}{E_{\text{ext}}} = a \ln \epsilon_{\text{max}} + b \tag{7.57}$$

where E_{DMRG} are the energies obtained from calculations with different m values and ϵ_{max} is the maximum truncation error from the corresponding calculation.

Instead of fixing m, one may start with a predefined truncation error ϵ_{max} and adjust m during every SVD truncation step (Eq. (7.47)) so that the truncation error does not exceed the predefined value. This approach, dubbed *dynamic block state selection* (DBSS) [60], allows to converge the energy up to a target accuracy. Instead of ϵ_{max}, it is also possible to have an entropy-based cutoff criterion [112], or to apply an iterative Richardson extrapolation scheme [61].

7.4.3 Post-DMRG Methods for Dynamic Correlation and Environment Effects

Although DMRG is capable of attaining an almost full-CI accuracy for active spaces up to 100 electrons in 100 orbitals in certain cases, this will usually not be sufficient to describe dynamic correlation in many large systems [113], where even thousands of orbitals may be correlated. Therefore, in the spirit of the traditional multi-configurational methods such as CASSCF, one may use DMRG-(SCF) to capture the static correlation of the system and complement it with an a-posteriori method to describe dynamic correlation. As for the CASSCF method, multi-configurational second-order perturbation-theory constitutes the first choice for the post-DMRG methods to describe dynamic correlation. Hence, implementations of the complete active space second-order perturbation theory and of the n-electron valence state second-order perturbation theory based on a DMRG reference wave function (DMRG-CASPT2 [114–117] and DMRG-NEVPT2 [118, 119]) have emerged: they are based on the reduced density matrix-based formulations of the perturbation theories and obtain the reduced density matrix elements from the DMRG wave function. This dependence on the reduced density matrices, in particular, higher-order reduced density matrices, turns out to be one of the largest drawbacks of these methods, especially when a reference wave function with a large active space is employed. For example, CASPT2 requires certain elements of the four-particle reduced density matrix (4e-RDM), while NEVPT2 requires full 4e-RDM, the evaluation of which scales as L^8. Due to this steep scaling, it is prohibitively expensive to evaluate the 4e-RDM for active space sizes with more than 30 orbitals. Higher-order RDMs may be approximated with the cumulant approximation [120–122]. The cumulant approximation for the RDMs has been implemented for NEVPT2 [123], DMRG-CASPT2 [116, 124] and also for the internally-contracted [125, 126] and the externally-contracted [127] multi-reference configuration interaction (DMRG-icMRCI and DMRG-ecMRCI, respectively). An

extensive benchmark study of cumulant-approximated DMRG-CASPT2 has been conducted by Phung et al. [128], stating that errors do not exceed 1–2 kcal mol^{-1} for active spaces with less than 24 active orbitals, but becomes more pronounced for larger active spaces. However, Zgid et al. [123] have shown that the cumulant approximation may introduce false intruder states in both CASPT2 and NEVPT2, as well as discontinuities in the potential energy surfaces. It is, therefore, even more welcome to see the development of the perturbation theory methods that avoid the higher-order RDMs altogether: the MPS perturbation theory by Sharma et al. [129, 130], which exploits the MPS structure and the sweep algorithm for both the reference and the first-order wave function; its combination with the internally-contracted NEVPT2 approach by the same authors [131]; the time-dependent formulation of NEVPT2 by Sokolov et al. [132, 133]; and the projected approximation by Roemelt et al. [134] Another method which does not rely on higher-order RDMs is the canonical transformation theory by Chan et al. [135–141], who employ an approximate similarity transformation of the Hamiltonian, transferring the complexity and the requirement to approximate the description of the dynamic correlation from the wave function to the Hamiltonian.

Density functional theory (DFT) has also been used in combination with DMRG to handle dynamic correlation. In particular, the short-range DFT, introduced by Savin et al. [142, 143] and applied in MCSCF calculations by Fromager et al. [144] separates the two-electron interaction into long-range and short-range parts. The short-range part is treated by DFT and the long-range part by wave function theory, e.g., a multi-configurational method, resulting in multi-configurational short-range DFT (MC-srDFT) [144–146]. We implemented an MC-srDFT variant with a DMRG reference wave function (DMRG-srDFT) [147]. Interestingly, DMRG-srDFT results are much less sensitive to the size of the active space, compared to standard CASSCF or DMRG calculations.

Veis et al. [148, 149] have combined the tailored coupled cluster (TCC) method with DMRG (DMRG-TCC). DMRG-TCC features a single-reference coupled cluster ansatz, where a part of the amplitudes are obtained from a preceeding DMRG calculation and are kept fixed in the CC procedure. For several multi-reference problems, the DMRG-TCCSD showed a better performance than standard single-reference CCSD. Another new development combining the coupled-cluster theory and DMRG is the multi-reference linearized coupled cluster theory based on MPS by Sharma and Alavi [150], which follows a similar ansatz to MPS-PT.

In addition to post-DMRG methods describing dynamic correlation, embedding schemes with DMRG have been developed. A DMRG-in-DFT embedding scheme [151] and an implementation of DMRG combined with polarizable embedding (PE) has recently been developed in our group [152].

Multi-configurational calculations on large systems, in addition to the requirement of larger active spaces, must also overcome the bottleneck of the increasing basis set size, and in particular, the bottleneck of handling the increasing number of two-electron integrals. Both well-established methods such as the resolution of the identity (RI) approximation [153–157] or the Cholesky decomposition of two-electron integrals [158–160], as well as novel domain-based pair natural orbital approximation [161] have found their way into multi-configurational methods and allow multi-configurational calculations to be performed on systems with over 5000 basis functions.

7.4.4 Analytical Energy Gradients and Non-Adiabatic Coupling Matrix Elements

Many phenomena, especially in photochemistry, rely on the optimization of excited state structures, conical intersections, potential energy surface crossings and ab initio molecular dynamic

simulations. They require not only the calculation of the ground-state and excited-state energies, but also of energy gradients and non-adiabatic couplings. Analytical energy gradients with respect to a perturbation x of a wave function which is fully variational with respect to all its parameters, may be calculated according to the Hellmann–Feynman theorem:

$$\frac{dE^{\Psi}}{dx} = \left\langle \Psi \left| \frac{\partial \hat{H}}{\partial x} \right| \Psi \right\rangle = \sum_{ij} \frac{\partial h_{ij}}{\partial x} \gamma_{ij} + \frac{1}{2} \sum_{ijkl} \frac{\partial V_{ijkl}}{\partial x} \Gamma_{ijkl}, \tag{7.58}$$

where $\partial h_{ij}/\partial x$ and $\partial V_{ijkl}/\partial x$ are the one- and two-electron integral derivatives, and γ_{ij} Γ_{ijkl} respectively are elements of the one- and two-particle RDMs. The former are of no concern to us since these are required in every analytical gradient implementation, and the latter may be obtained from a converged DMRG wave function with Eq. (7.30). However, the DMRG wave function is fully variational with respect to all of its parameters only in the case of a single-state DMRG-SCF wave function, i.e., when both the orbitals and the MPS are variationally optimized. If this is not the case, the RDMs in Eq. (7.58) must be replaced with the so-called *effective* RDMs $\boldsymbol{\gamma}^{\mathrm{e}}$ and $\boldsymbol{\Gamma}^{\mathrm{e}}$, respectively,

$$\boldsymbol{\gamma}^{\mathrm{e}} = \boldsymbol{\gamma} + \boldsymbol{\gamma}^{\mathrm{o}} + \boldsymbol{\gamma}^{\mathrm{MPS}}, \tag{7.59}$$

$$\boldsymbol{\Gamma}^{\mathrm{e}} = \boldsymbol{\Gamma} + \boldsymbol{\Gamma}^{\mathrm{o}} + \boldsymbol{\Gamma}^{\mathrm{MPS}}, \tag{7.60}$$

where the matrices with the superscripts $^{\mathrm{o}}$ and $^{\mathrm{MPS}}$ denote the contributions from the non-variationality of the wave function with respect to the orbitals and the MPS, respectively. These can be calculated with the help of a Lagrangian [162–164], and the corresponding Lagrange multipliers are obtained from the solution of the *coupled-perturbed* MCSCF equations [165, 166], whose form is identical to Eq. (6.156) in Chapter 6 of this book:

$$\begin{pmatrix} {}^{\mathrm{cc}}\mathbf{E}^{(2)} & {}^{\mathrm{oc}}\mathbf{E}^{(2)} \\ {}^{\mathrm{co}}\mathbf{E}^{(2)} & {}^{\mathrm{oo}}\mathbf{E}^{(2)} \end{pmatrix} \begin{pmatrix} \mathbf{S} \\ \mathbf{R} \end{pmatrix} = - \begin{pmatrix} {}^{\mathrm{c}}\mathbf{E}^{(1)} \\ {}^{\mathrm{o}}\mathbf{E}^{(1)} \end{pmatrix}. \tag{7.61}$$

For energy gradients, ${}^{\mathrm{c}}\mathbf{E}^{(1)}$ is zero, $\mathbf{R}$ are orbital Lagrange multipliers, and for DMRG wave functions $\mathbf{S}$ are MPS Lagrange multipliers. For a state-specific DMRG (not DMRG-SCF) gradient, the MPS contribution in Eqs. (7.59) and (7.60), as well as $\mathbf{S}$ in Eq. (7.61) vanish, and Eq. (7.61) must be solved only for the orbital block. Hu and Chan [167] first presented such an implementation of analytical DMRG gradients.

In the case of a state-averaged DMRG-SCF wave function (cf. Chapter 6), the wave function is no longer variational, neither with respect to orbitals nor to the MPS, and hence Eq. (7.61) must be solved in its full form. This raises the question of the parameter space, or, in other words, which basis defines $\mathbf{S}$? For comparison, in state-averaged CASSCF analytical gradients $\mathbf{S}$ is defined in the basis of configuration state functions (CSFs), which is not possible for a DMRG wave function. In our recent work [168], we addressed this problem and presented an implementation of approximate state-averaged DMRG-SCF analytical gradients. Rather than employing the space encompassing all MPS tensors (which would be very large and highly redundant), we may choose a site (preferably in the middle of the lattice), and express the MPS at the mixed-canonical form at this site, analogously to Eq. (7.52):

$$|\Psi\rangle = \sum_{\sigma_l a_{l-1}, a_l} M^{\sigma_l}_{a_{l-1} a_l} |a_{l-1}\rangle \otimes |\sigma_l\rangle \otimes |b_l\rangle. \tag{7.62}$$

The MPS Lagrange multipliers $\mathbf{S}$ may now be defined in the basis $|a_{l-1}\rangle \otimes |\sigma_l\rangle \otimes |b_l\rangle$ and in the same parameter space as the MPS tensor $M^{\sigma_l}_{a_{l-1} a_l}$. With this definition we may obtain expressions

for **S** similar to those found in the CASSCF analytical gradient formulation and similarly solve Eq. (7.61). An expression for the non-adiabatic coupling matrix elements between states is obtained in a similar manner to gradients [169–171]: the main difference is that transition density matrices between the states in question instead of state-specific density matrices are employed in Eqs. (7.59) and (7.60).

7.4.5 Tensor Network States

The one-dimensional structure of the DMRG lattice and the MPS is naturally best for describing one-dimensional systems, or at least systems, in which the correlation is between neighboring sites. In fact, the most accurate and the largest calculations have been performed for one-dimensional or quasi one-dimensional systems [63]. (see Section 7.5) Nevertheless, DMRG still works well in describing the correlation in molecules (which is in general, three-dimensional), but in general usually requires a much larger m value to obtain an accurate solution. One would, however, expect a higher-dimensional lattice to capture the multi-dimensional correlation better, and therefore, several multi-dimensional generalizations of MPS have been designed – *tensor network states* (TNS) with various topologies.

Many TNS topologies have been developed (see e.g., Refs. [172–175] and references therein) and have found applications in physics, however, only two types have been applied in quantum chemistry so far [176]: the complete graph tensor network states (CGTNS) [177–179], where in a lattice every site is connected to all other sites and the tree tensor network states (TTNS) [108, 180–183] featuring a tree topology. Figures 7.5(b) and (c) show topologies of example CGTNS and TTNS, respectively.

The more complicated topology of TNS, while being its largest advantage when it comes to capturing correlation, is also its largest disadvantage: the optimization of the parameters of a high-dimensional tensor network is highly non-trivial and is of much higher numerical cost. Many TNS, in particular CGTNS cannot be optimized variationally and one must resort to numerical techniques such as Monte-Carlo optimization [178]. Although TTNS allow DMRG-like variational optimization algorithms [183] the scaling is still worse than that of traditional DMRG. For these reasons, TNS in quantum chemistry have not yet been established and the progress in the field of TNS in recent years has been fairly limited.

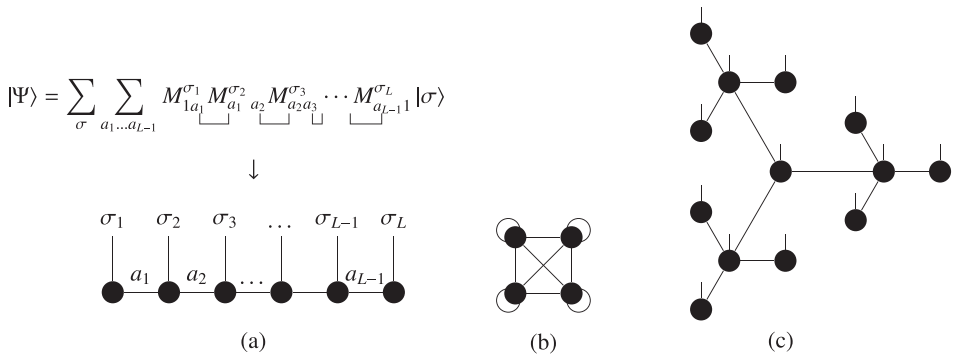


Figure 7.5 (a) Graphical representation of an MPS. The circles denote tensors, the *virtual* indices a_i are contracted over and are denoted by the connecting lines, the *physical* indices σ_i are denoted with the unconnected lines. (b) An analogous representation of an example topology of CGTNS and (c) for TTNS.

7.5 Applications in Quantum Chemistry

DMRG and post-DMRG methods have become powerful tools in the multi-configurational toolbox. The majority of DMRG applications can be divided into three classes. The first class constitutes the classical benchmark applications, which are primarily used to demonstrate the applicability and the performance of the methods. Typical benchmark studies include ground state energies of ethylene and higher polyenes [47, 51, 63, 125, 138, 184, 185], hydrogen chains, rings or two-dimensional aggregates [46, 47, 56, 57, 63, 186–188], water [42, 47–49, 60, 61, 88, 104, 112, 141, 152, 188, 189], dissociation curves of N_2 [44, 45, 47, 50, 88, 89, 91, 107, 125, 132, 141, 148, 187, 189], the chromium dimer [44, 53, 58, 75, 107, 110, 114, 115, 118, 132, 148, 150, 187, 189] and relative energies of the bis-μ-oxo and the μ-η^2 : η^2-peroxo isomers of Cu_2O_2 [58, 108, 138, 190].

The second class exploits the one-dimensional nature of DMRG and comprises studies of strongly correlated problems in pseudo-one-dimensional systems. Examples of applications belonging to this class are studies of the radical character of polyacenes [184], polyphenylenecarbenes [191, 192], but also graphene nanoribbons [185] or meta-benzyne [193]. There are also studies of excited state properties of polyenes [104, 109, 167] and other elongated systems. Most commonly, in these applications DMRG is used as a standalone method, although post-DMRG methods such as DMRG-CASPT2 are occasionally used as well [194]. Furthermore, many polyene studies employ the Pariser–Parr–Pople (PPP) Hamiltonian [40, 41, 195–197], which has also been employed to study polyphenylenes [198], polyacenes [199], polythiophenes [200, 201], phenyl- and stilbene-substituted dendrimers [202], graphene nanoribbons [203], and also non-elongated systems such as porphines and metalloporphines [204, 205].

The final (and, by far, the largest) class of DMRG applications exploits its capability as an approximate CI solver and constitutes applications to a broad range of multi-reference problems which require a large active space. In these studies DMRG is generally employed with orbital optimization (i.e., as the DMRG-SCF method) and is complemented by post-DMRG methods describing dynamic correlation. Below we discuss some application examples from this class, sorted into different categories.

Bioinorganic Chemistry

A large number of studies focus on the electronic structure of transition metal complexes and clusters, the majority of which find their application in bioinorganic chemistry. Although DMRG should not be applicable to such problems due to its inherent one-dimensional structure, we argued 10 years ago [190] that it still may be superior to traditional approaches.

Biomimetic manganese complexes that mimic the manganese cluster in photosystem II [206] and find application as water oxidation catalysts [207, 208] have received much attention in studies with DMRG, due to the strong multi-reference nature of the Mn cluster and its requirement for a large active space. Binding of first-row diatomics to heme models has been also shown to be a highly multi-reference problem and a challenging computational task requiring large basis sets and high-level correlation treatment [209], hence Phung and Pierloot [210] very recently performed a study on electronic structure and binding energies of Fe and Mn porphyrin-O_2 adducts. Metal-nitrosyl bonds, like metal-O_2 bonds show an intricate electronic structure and a strong multi-reference character (NO is a well-known non-innocent ligand): several studies of metal nitrosyl complexes have been performed in our group in the past [95, 119, 211]. Other bioinorganic chemistry-related studies include reaction mechanism studies on the non-heme iron active site of Δ^9 desaturase [212], diferrate-mediated water oxidation [213], η^2 metal-ethene adduct formation [214] or on H_2 binding to the active site of the NiFe hydrogenase [215]. In addition to the

electronic structure and energies, spin densities [211, 216] and magnetic couplings [208, 217–221] are also of interest. A special mention is warranted of studies of spin-state energetics of spin crossover compounds [222], which will be described next.

Spin Crossover Compounds

Spin symmetry plays an important role in the electronic structure and reactivity of transition metal complexes. Many transition metal complexes (dubbed *spin crossover* (SCO) compounds) [223–227] have close-lying states of different spin multiplicities and may undergo spin crossover even upon heating. Even if a transition metal complex does not show spin crossover behavior, different spin states may still be important for its reactivity [228, 229]. Hence, the correct description of the spin and electronic structure of the ground and the lower excited states is crucial for the chemistry of transition metal complexes. It is, however, equally challenging [225–227], given the large multi-reference character in many transition metal complexes, the energetic vicinity of different spin states in SCO compounds and the resulting high demands to the accuracy of theoretical methods to even qualitatively predict the correct spin state. Unsurprisingly, there have been many spin-state studies with DMRG and post-DMRG methods [104, 119, 126, 128, 190, 215, 222, 230–233]. As an example, below we will present a study on electronic structure and spin-state energetics of a cobalt tropocoronand complex with DMRG-SCF and DMRG-NEVPT2, carried out in our group [119].

Tropocoronand ligands (see Figure 7.6(a)) provide a porphyrin-like constrained coordination environment for metal complexes, tunable by varying length n of the alkyl chain. Transition metal nitrosyl complexes show intricate electronic structure due to the non-innocence of the NO ligand and cannot be reliably described by single-reference methods such as density functional theory [95]. Franz et al. [234] synthesized and characterized cobalt tropocoronand nitrosyl complexes with $n = 3$ and 4. They found that the complex with $n = 3$ ([Co(TC-3,3)NO], Figure 7.6(b) was paramagnetic, which, however, was in contradiction with a later DFT and experimental reinvestigation [235]. Based on this controversy and the multi-reference character of the complex, we opted for a reinvestigation of the electronic structure and the spin-state energetics of [Co(TC-3,3)NO] with DMRG-SCF and DMRG-NEVPT2 employing Cholesky decomposition of the two-electron integrals (CD-DMRG-NEVPT2).

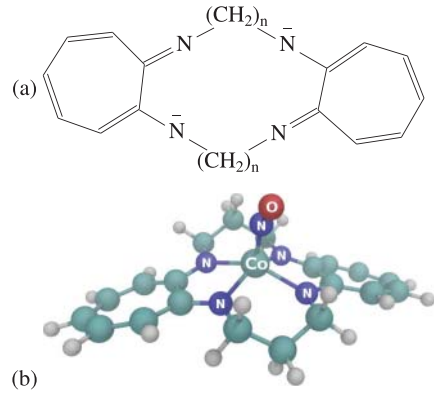


Figure 7.6 (a) The tropocoronand ligand; (b) [Co(TC-3,3)NO]. Adapted from Ref. [119].

Table 7.2 Singlet-triplet energy gap (in kcal mol^{-1}) of [Co(TC-3,3)(NO)] calculated with CD-DMRG-NEVPT2 (abbreviated as NEVPT2) and other methods. Adapted from Ref. [119]. DFT results from Ref. [235] are provided for the state of an equivalent character.

	NEVPT2	DMRG-SCF	OLYP [235]	PW91 [235]	B3LYP-D3 [235]
T_1	35.0	38.6	–	–	–
T_2	36.1	29.6	23.8	25.1	10.4

The S_0 state shows the natural orbital occupation numbers (NOON) for the $3d_{z^2}$ and the $\pi^*_{\mathrm{NO},y}$ orbitals strongly deviating from 2 or 0, which is a sign of a strong multi-reference character typical for metal nitrosyl complexes [95]: also the T_1 state shows NOON for the $3d_{yz}$ and the $\pi^*_{\mathrm{NO},y}$ close to 1.5 and 0.5, which corresponds to a situation between a d^7 and d^8 cobalt and neutral and cationic NO. This strongly differs from the state character described by the DFT calculation in the previous study by Hopmann et al. [235], which characterizes T_1 state as d^7 Co and neutral NO. Also the character of T_2 in our work differs from that in Ref. [235], even though it hardly shows any multi-configurational character according to NOONs. The single-configurational nature of the Kohn–Sham DFT does not allow for an adequate description of multi-reference states, and therefore the discrepancies between DFT and our DMRG-SCF state characters are not surprising.

Having characterized the electronic structure of the states in [Co(TC-3,3)(NO)] with DMRG-SCF, we also calculated the S_0–T_1 and S_0–T_2 energy gaps with CD-DMRG-NEVPT2, which are presented in Table 7.2. Although NEVPT2 has been presented in two different formulations in the original literature [236–238] – the "strongly contracted" and the "partially contracted" formulation, we have found that the "partially contracted" formulation shows numerical instabilities when used with a DMRG reference wave function obtained with small m values, whereas the "strongly contracted" formulation does not show this behavior. Therefore, all NEVPT2 results reported below are obtained with the "strongly contracted" formulation.

Both CD-DMRG-NEVPT2 and DMRG-SCF predict S_0 as the ground state, agreeing with the new experimental data and all DFT functionals in the study by Hopmann et al. [235], contradicting the previous reports of the paramagneticity of the complex. However, the CD-DMRG-NEVPT2 singlet–triplet gaps are at least 10 kcal mol^{-1} larger than the results obtained with DFT.

Heavy Element Chemistry

Heavy element studies require an adequate treatment of relativistic effects. Both four-component DMRG [220, 239] and the simplified treatment of scalar relativistic effects with the Douglas–Kroll–Hess transformation [54, 218] and the spin–orbit coupling with the state-interaction method [219, 221, 240, 241] have been implemented. We have formulated the state-interaction approach for arbitrary MPS with non-orthogonal MO bases [219], named MPS-SI. The key concept behind the method is the transformation of the MPS into a biorthonormal basis representation [242], similar to the state interaction approach by Malmqvist and Roos [243] for CI wave functions. Heavy element systems studied by DMRG include CsH [54, 91], TlH [220, 239], plutonium oxides [219, 244], a dysprosium(III) complex [220] or noble-gas coordination reaction to CUO [245].

Photochemistry

Photochemical studies involving DMRG are also abundant. While we have mentioned some prototypical studies of polyenes and other quasi-one-dimensional systems at the beginning of this section, there are certainly more. Some studies focus on potential energy curves of diatomics: LiF

[52], CsH [54] and C_2 [246]; others on elucidation of photophysical and photochemical processes such as the ring opening of spiropyran [247], singlet fission in donor–acceptor copolymers [248], vinylogous Nazarov-type photocyclization [249], the delayed fluorescence in carbene-metal amides [250], or the electronic structure of a naphtalene excimer [251]. State-specific analytical ground- and excited-state gradients for DMRG-SCF wave functions have been developed, allowing for excited state structure optimizations [167] or the calculation of resonance Raman spectra [252]. Recently, we have also developed a formulation for analytical gradients for state-averaged DMRG-SCF wave functions [168]. Below we will discuss the resonance Raman spectra of uracil from Ref. [252] in more detail.

Resonance Raman (RR) spectroscopy [253, 254] allows the selective observation of vibrational modes that are associated with a selected electronic transition, providing access to the structure and dynamics of a particular excited state. Hence, calculations of RR spectra are required for the interpretation of the experimental studies, enhance our understanding of the RR effect and provide insights into the nature of the excited state in question. Investigating excited states of DNA nucleobases gives insights into DNA photostability and photodamage mechanisms [255].

By following the Kramers–Kronig relation [256] and several simplifications, including (a) neglecting vibronic couplings, (b) considering only one excited state in resonance, (c) independent mode displaced harmonic oscillator (IMDHO) model [31–35]), the RR intensities may be calculated employing the ground state vibrational frequencies, the gradients of the excited state energy with respect to ground state normal coordinates at the ground-state equilibrium position and the transition dipole moment between the ground and the excited state (for details, see Refs. [252] and [254] and references therein).

Previous studies of RR spectra of uracil [254, 257–259] showed the need to obtain an accurate excited state PES and excited state gradients. For this reason, the RR spectra of uracil in Ref. [252] have been calculated with TD-DFT (PBE0), CI singles, CASSCF, CASSCF/CASPT2 and DMRG-SCF employing various active spaces. An incident wavelength of 266 nm and an absorption band maximum of 234 nm (corresponding to 5.29 eV, i.e., the excitation energy for the S_2 obtained with TD-DFT) were chosen for all calculations.

First, the spectra were calculated with CASSCF and CASPT2 employing an active space comprising 14 electrons in 10 orbitals (14,10), as in the study of Mercier et al. [260] However, a large discrepancy between the CASSCF and the CASPT2 RR intensities and the S_2 excitation energies (but not S_1 excitation energies) was found, which led to the belief that this active space does not allow for an exhaustive description of the static correlation effects in the S_2 state. Therefore, several larger active spaces have been devised: (14,18) and (14,26), corresponding to the (14,10) active space augmented with different sets of Rydberg orbitals, (30,26), corresponding to the full valence space of uracil (but without Rydberg orbitals) and (18,23), a reduced valence space retaining orbitals in the bonds important for the structural changes during the excitation and the subsequent structure relaxation in the S_2 state.

Figure 7.7 shows the RR spectra for uracil calculated with different methods. In general the TD-DFT/PBE0 spectrum reproduces the experimental spectrum fairly well, except for the relative intensities of the peaks above 1400 cm^{-1}. Of the multi-configurational methods, DMRG-SCF(30,26), not surprisingly, yields best results. The largest discrepancies with the experiment are found in the low-frequency modes, which are delocalized ring bending/stretching modes (see Ref. [260] for the peak assignment). In particular, the correct peak intensity ratio of the two most intensive low-frequency modes at 579 and 789 cm^{-1} is reproduced only with DMRG-SCF(30,26) and with PBE0, indicating that only the full-valence (30,26) active space provides a balanced

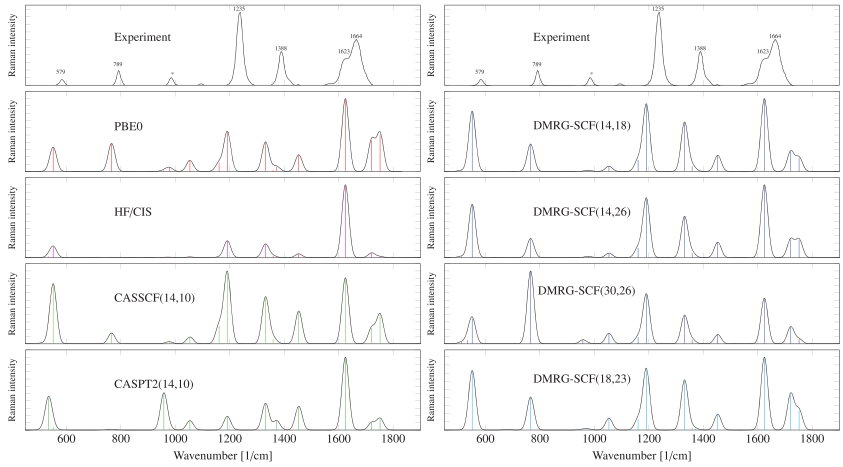


Figure 7.7 Resonance Raman spectra for uracil, calculated with different methods. Adapted from figures 4 and 5 of Ref. [252].

description of the electron correlation. Concerning the effect of Rydberg orbitals on the RR intensities, the addition of Rydberg orbitals balances the intensity ratio of the two highest-frequency peaks which correspond to the carbonyl stretches.

In summary, of the multi-configurational methods, DMRG-SCF(30,26) was found to give the best qualitative agreement between the calculated and the experimental spectrum. DMRG-SCF(18,23) spectrum agrees well with DMRG-SCF(30,26), with a notable exception for the intensities of the low-frequency delocalized modes. These results, therefore, indicate the importance of a choice of a balanced active space for the multi-configurational calculation.

Vibrational DMRG

In addition to the electronic structure problems in quantum chemistry, vibrational DMRG (vDMRG) has been developed in our group [261]. vDMRG is based on a vibrational CI ansatz, where a full CI expansion is carried out in a basis of the harmonic vibrational modes. In analogy to the full CI of an electronic problem, vDMRG approximates the full vCI solution. As the Hamiltonian, an approximate Watson Hamiltonian [262] (i.e., a form of a second-quantized vibrational Hamiltonian) is employed: the flexibility of the MPS-MPO ansatz allows the construction of the Watson Hamiltonian as an MPO and performing a subsequent DMRG calculation including the same formalism (and the implementation) of the DMRG algorithm as for the electronic problem. vDMRG also closely relates to the vibrational tensor-train approach of Rakhuba and Oseledets [263].

DMRG Software

Currently, there are five widely-used computer programs available for DMRG calculations in quantum chemistry: QC-DMRG-Budapest by Legeza et al. [60], the program by Kurashige and Yanai [58], Block by the Chan group [47, 49, 71, 75, 109], CheMPS2 by Wouters et al. [246] and QCMaquis developed in our group [81, 264]. The last three programs are available for free: Block from Ref. [265], CheMPS2 from Ref. [266] and QCMaquis from Ref. [267].

All of the DMRG programs require precalculated two-electron integrals and information about the active orbital space. To obtain these, a general-purpose quantum chemistry programs is usually involved, or the DMRG calculation is performed employing an interface between a DMRG code and a general-purpose program.

Block and CheMPS2 implement the DMRG algorithm in its traditional renormalization group formulation, although they use MPS to represent the wave function. QCMaquis, on the other hand, presents a DMRG implementation based fully on the MPS-MPO formalism. Other MPS-MPO implementations have been presented by Nakatani [268] and by Legeza, however they have not been described in detail in the literature.

7.6 Conclusions

In this chapter, we attempted to give the reader an overview of the concepts, theory, the new developments in density matrix renormalization group and its applications in quantum chemistry. Although originally designed for one-dimensional or quasi-one-dimensional systems, at its best, DMRG has found applications in many other quantum chemical systems. Together with the post-DMRG methods it has found its place within the multi-configurational quantum chemistry methods for large active spaces. The orbital entanglement measures have proven themselves as a

very valuable concept for the analysis of electron correlation and the automated active space selection. Although in recent years many new methods to solve the CI problem for large active spaces have been developed, potentially being competitors for DMRG, the popularity of DMRG-based methods is still on the rise.

Acknowledgment

L.F. acknowledges the Austrian Science Fund for a Schrödinger fellowship (Project No. J 3935).

References

1 Roos, B.O., Taylor, P.R., and Siegbahn, P.E. (1980). A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach. *Chem. Phys.* 48: 157–173.

2 Roos, B.O. (1980). The complete active space SCF method in a Fock-matrix-based super-CI formulation. *Int. J. Quantum Chem.* 18: 175–189.

3 Cheung, L.M., Sundberg, K.R., and Ruedenberg, K. (1978). Dimerization of carbene to ethylene. *J. Am. Chem. Soc.* 100: 8024–8025.

4 Cheung, L.M., Sundberg, K.R., and Ruedenberg, K. (1979). Electronic rearrangements during chemical reactions. II. Planar dissociation of ethylene. *Int. J. Quantum Chem.* 16: 1103–1139.

5 Ruedenberg, K., Schmidt, M.W., Gilbert, M.M., and Elbert, S.T. (1982). Are atoms intrinsic to molecular electronic wave functions? I. The FORS model. *Chem. Phys.* 71: 41–49.

6 Vogiatzis, K.D., Ma, D., Olsen, J. et al. (2017). Pushing configuration-interaction to the limit: towards massively parallel MCSCF calculations. *J. Chem. Phys.* 147: 184111.

7 Malmqvist, P.A., Rendell, A., and Roos, B.O. (1990). The restricted active space self-consistent-field method, implemented with a split graph unitary group approach. *J. Phys. Chem.* 94: 5477–5482.

8 Ma, D., Li Manni, G., and Gagliardi, L. (2011). The generalized active space concept in multiconfigurational self-consistent field methods. *J. Chem. Phys.* 135: 044128.

9 Ivanic, J. (2003). Direct configuration interaction and multiconfigurational self-consistent-field method for multiple active spaces with variable occupations. I. Method. *J. Chem. Phys.* 119: 9364–9376.

10 Li Manni, G., Aquilante, F., and Gagliardi, L. (2011). Strong correlation treated via effective hamiltonians and perturbation theory. *J. Chem. Phys.* 134: 034114.

11 Li Manni, G., Ma, D., Aquilante, F. et al. (2013). SplitGAS method for strong correlation and the challenging case of Cr_2. *J. Chem. Theory Comput.* 9: 3375–3384.

12 Zimmerman, P.M. (2017). Incremental full configuration interaction. *J. Chem. Phys.* 146: 104102.

13 Eriksen, J.J., Lipparini, F., and Gauss, J. (2017). Virtual orbital many-body expansions: a possible route towards the full configuration interaction limit. *J. Phys. Chem. Lett.* 8: 4633–4639.

14 Eriksen, J., and Gauss, J. (2018). Many-body expanded full configuration interaction. I. Weakly correlated regime. *J. Chem. Theory Comput.* 14: 5180–5191.

15 Booth, G.H., Thom, A.J.W., and Alavi, A. (2009). Fermion Monte Carlo without fixed nodes: a game of life, death, and annihilation in Slater determinant space. *J. Chem. Phys.* 131: 054106.

16 Li Manni, G., Smart, S.D., and Alavi, A. (2016). Combining the complete active space self-consistent field method and the full configuration interaction quantum monte carlo within a super-CI framework, with application to challenging metal-porphyrins. *J. Chem. Theory Comput.* 12: 1245–1258.

17 Holmes, A.A., Tubman, N.M., and Umrigar, C.J. (2016). Heat-bath configuration interaction: an efficient selected configuration interaction algorithm inspired by heat-bath sampling. *J. Chem. Theory Comput.* 12: 3674–3680.

18 Smith, J.E.T., Mussard, B., Holmes, A.A., and Sharma, S. (2017). Cheap and near exact CASSCF with large active spaces. *J. Chem. Theory Comput.* 13: 5468–5478.

19 Tubman, N.M., Lee, J., Takeshita, T.Y. et al. (2016). A deterministic alternative to the full configuration interaction quantum monte carlo method. *J. Chem. Phys.* 145: 044112.

20 Tubman, N.M., Freeman, C.D., Levine, D.S. et al. (2020). Modern approaches to exact diagonalization and selected configuration interaction with the adaptive sampling CI method. *J. Chem. Theory Comput.* 16: 2139–2159.

21 Petruzielo, F.R., Holmes, A.A., Changlani, H.J. et al. (2012). Semistochastic projector Monte Carlo method. *Phys. Rev. Lett.* 109: 230201.

22 Holmes, A.A., Changlani, H.J., and Umrigar, C.J. (2016). Efficient heat-bath sampling in Fock space. *J. Chem. Theory Comput.* 12: 1561–1571.

23 Huron, B., Malrieu, J.P., and Rancurel, P. (1973). Iterative perturbation calculations of ground and excited state energies from multiconfigurational zeroth-order wave functions. *J. Chem. Phys.* 58: 5745–5759.

24 Buenker, R.J. and Peyerimhoff, S.D. (1974). Individualized configuration selection in CI calculations with subsequent energy extrapolation. *Theoret. Chim. Acta* 35: 33–58.

25 Miralles, J., Daudey, J.-P., and Caballol, R. (1992). Variational calculation of small energy differences. The singlet-triplet gap in $[Cu_2Cl_6]^{2-}$. *Chem. Phys. Lett.* 198: 555–562.

26 Miralles, J., Castell, O., Caballol, R., and Malrieu, J.-P. (1993). Specific CI calculation of energy differences: transition energies and bond energies. *Chem. Phys.* 172: 33–43.

27 Neese, F. (2003). A spectroscopy oriented configuration interaction procedure. *J. Chem. Phys.* 119: 9428–9443.

28 Olsen, J., Malmqvist, P.-A., Roos, B.O. et al. (1987). A non-linear approach to configuration interaction: the low-rank CI method (LR CI). *Chem. Phys. Lett.* 133: 91–101.

29 Koch, H. and Dalgaard, E. (1992). A variational matrix decomposition applied to full configuration-interaction calculations. *Chem. Phys. Lett.* 198: 51–58.

30 Fales, B.S., Seritan, S., Settje, N.F. et al. (2018). Large-scale electron correlation calculations: rank-reduced full configuration interaction. *J. Chem. Theory Comput.* 14: 4139–4150.

31 Shepard, R. (2005). A general nonlinear expansion form for electronic wave functions. *J. Phys. Chem. A* 109: 11629–11641.

32 Shepard, R. and Minkoff, M. (2006). Optimization of nonlinear wave function parameters. *Int. J. Quantum Chem.* 106: 3190–3207.

33 Shepard, R., Minkoff, M., and Brozell, S.R. (2007). Nonlinear wave function expansions: a progress report. *Int. J. Quantum Chem.* 107: 3203–3218.

34 Gidofalvi, G. and Shepard, R. (2009). The evaluation of spin-density matrices within the graphically contracted function method. *Int. J. Quantum Chem.* 109: 3552–3563.

35 Gidofalvi, G. and Shepard, R. (2009). Computation of determinant expansion coefficients within the graphically contracted function method. *J. Comput. Chem.* 30: 2414–2419.

36 Shavitt, I. (1977). Graph theoretical concepts for the unitary group approach to the many-electron correlation problem. *Int. J. Quantum Chem.* 12: 131–148.

37 Shavitt, I. (1978). Matrix element evaluation in the unitary group approach to the electron correlation problem. *Int. J. Quantum Chem.* 14: 5–32.

38 White, S.R. (1992). Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.* 69: 2863–2866.

39 White, S.R. (1993). Density-matrix algorithms for quantum renormalization groups. *Phys. Rev. B* 48: 10345–10356.

40 Fano, G., Ortolani, F., and Ziosi, L. (1998). The density matrix renormalization group method: application to the PPP model of a cyclic polyene chain. *J. Chem. Phys.* 108: 9246–9252.

41 Yaron, D., Moore, E.E., Shuai, Z., and Brédas, J.L. (1998). Comparison of density matrix renormalization group calculations with electron-hole models of exciton binding in conjugated polymers. *J. Chem. Phys.* 108: 7451–7458.

42 White, S.R. and Martin, R.L. (1999). Ab initio quantum chemistry using the density matrix renormalization group. *J. Chem. Phys.* 110: 4127–4130.

43 Daul, S., Ciofini, I., Daul, C., and White, S.R. (2000). Full-CI quantum chemistry using the density matrix renormalization group. *Int. J. Quantum Chem.* 79: 331–342.

44 Mitrushenkov, A.O., Fano, G., Ortolani, F. et al. (2001). Quantum chemistry using the density matrix renormalization group. *J. Chem. Phys.* 115: 6815–6821.

45 Mitrushenkov, A.O., Linguerri, R., Palmieri, P., and Fano, G. (2003). Quantum chemistry using the density matrix renormalization group II. *J. Chem. Phys.* 119: 4148–4158.

46 Mitrushchenkov, A.O., Fano, G., Linguerri, R., and Palmieri, P. (2011). On the importance of orbital localization in QC-DMRG calculations. *Int. J. Quantum Chem.* 112: 1606–1619.

47 Chan, G.K.-L. and Head-Gordon, M. (2002). Highly correlated calculations with a polynomial cost algorithm: a study of the density matrix renormalization group. *J. Chem. Phys.* 116: 4462–4476.

48 Chan, G.K.-L. and Head-Gordon, M. (2003). Exact solution (within a triple-zeta, double polarization basis set) of the electronic schrödinger equation for water. *J. Chem. Phys.* 118: 8551–8554.

49 Chan, G.K.-L. (2004). An algorithm for large scale density matrix renormalization group calculations. *J. Chem. Phys.* 120: 3172–3178.

50 Chan, G.K.-L., Kállay, M., and Gauss, J. (2004). State-of-the-art density matrix renormalization group and coupled cluster theory studies of the nitrogen binding curve. *J. Chem. Phys.* 121: 6110–6116.

51 Chan, G.K.-L. and Van Voorhis, T. (2005). Density-matrix renormalization-group algorithms with nonorthogonal orbitals and non-hermitian operators, and applications to polyenes. *J. Chem. Phys.* 122: 204101.

52 Legeza, Ö., Röder, J., and Hess, B.A. (2003). QC-DMRG study of the ionic-neutral curve crossing of LiF. *Mol. Phys.* 101: 2019–2028.

53 Moritz, G., Hess, B.A., and Reiher, M. (2005). Convergence behavior of the density-matrix renormalization group algorithm for optimized orbital orderings. *J. Chem. Phys.* 122: 024107.

54 Moritz, G., Wolf, A., and Reiher, M. (2005). Relativistic DMRG calculations on the curve crossing of cesium hydride. *J. Chem. Phys.* 123: 184105.

55 Zgid, D. and Nooijen, M. (2008). On the spin and symmetry adaptation of the density matrix renormalization group method. *J. Chem. Phys.* 128: 014107.

56 Zgid, D. and Nooijen, M. (2008). The density matrix renormalization group self-consistent field method: orbital optimization with the density matrix renormalization group method in the active space. *J. Chem. Phys.* 128: 144116.

57 Zgid, D. and Nooijen, M. (2008). Obtaining the two-body density matrix in the density matrix renormalization group method. *J. Chem. Phys.* 128: 144115.

58 Kurashige, Y. and Yanai, T. (2009). High-performance ab initio density matrix renormalization group method: applicability to large-scale multireference problems for metal compounds. *J. Chem. Phys.* 130: 234114.

59 Legeza, O. and Fáth, G. (1996). Accuracy of the density-matrix renormalization-group method. *Phys. Rev. B* 53: 14349–14358.

60 Legeza, O., Röder, J., and Hess, B.A. (2003). Controlling the accuracy of the density-matrix renormalization-group method: the dynamical block state selection approach. *Phys. Rev. B* 67: 125114.

61 Marti, K.H. and Reiher, M. (2010). DMRG control using an automated Richardson-type error protocol. *Mol. Phys.* 108: 501–512.

62 Hubig, C., Haegeman, J., and Schollwöck, U. (2018). Error estimates for extrapolations with matrix-product states. *Phys. Rev. B* 97: 045125.

63 Hachmann, J., Cardoen, W., and Chan, G.K.-L. (2006). Multireference correlation in long molecules with the quadratic scaling density matrix renormalization group. *J. Chem. Phys.* 125: 144101.

64 Schollwöck, U. (2011). The density-matrix renormalization group in the age of matrix product states. *Ann. Phys.* 326: 96–192.

65 Schollwöck, U. (2011). The density-matrix renormalization group: a short introduction. *Phil. Trans. R. Soc. A* 369: 2643–2661.

66 Hallberg, K. (2004). *Theoretical Methods for Strongly Correlated Electrons* (eds. D. Sénéchal, A.-M. Tremblay and C. Bourbonnais), 3–37. Springer.

67 Hallberg, K.A. (2006). New trends in density matrix renormalization. *Adv. Phys.* 55: 477–526.

68 Chan, G.K.-L., Dorando, J.J., Ghosh, D. et al. (2008). *Frontiers in Quantum Systems in Chemistry and Physics*, Progress in Theoretical Chemistry and Physics 18 (eds. S. Wilson, P.J. Grout, J. Maruani, et al.), 49–65. Springer Netherlands.

69 Chan, G.K.-L. and Zgid, D. (2009). *Annual Reports in Computational Chemistry*, vol. 5 (ed. R.A. Wheeler), 149–162. Elsevier.

70 Chan, G.K.-L. and Sharma, S. (2011). The density matrix renormalization group in quantum chemistry. *Annu. Rev. Phys. Chem.* 62: 465–481.

71 Olivares-Amaya, R., Hu, W., Nakatani, N. et al. (2015). The ab-initio density matrix renormalization group in practice. *J. Chem. Phys.* 142: 034102.

72 Chan, G.K.-L., Keselman, A., Nakatani, N. et al. (2016). Matrix product operators, matrix product states, and ab initio density matrix renormalization group algorithms. *J. Chem. Phys.* 145: 014102.

73 Marti, K.H. and Reiher, M. (2011). New electron correlation theories for transition metal chemistry. *Phys. Chem. Chem. Phys.* 13: 6750–6759.

74 Wouters, S. and Van Neck, D. (2014). The density matrix renormalization group for ab initio quantum chemistry. *Eur. Phys. J. D* 68: 272.

75 Sharma, S. and Chan, G.K.-L. (2012). Spin-adapted density matrix renormalization group algorithms for quantum chemistry. *J. Chem. Phys.* 136: 124121.

76 Keller, S. and Reiher, M. (2016). Spin-adapted matrix product states and operators. *J. Chem. Phys.* 144: 134101.

77 Jordan, P. and Wigner, E. (1928). Über das Paulische Äquivalenzverbot. *Z. Physik* 47: 631–651.

78 Freitag, L. (2015). Quantum chemical studies on electronic structure and photodynamics of ruthenium complexes. Ph.D. thesis, University of Vienna.

79 Östlund, S. and Rommer, S. (1995). Thermodynamic limit of density matrix renormalization. *Phys. Rev. Lett.* 75: 3537–3540.

80 Rommer, S. and Östlund, S. (1997). Class of ansatz wave functions for one-dimensional spin systems and their relation to the density matrix renormalization group. *Phys. Rev. B* 55: 2164–2181.

81 Keller, S., Dolfi, M., Troyer, M., and Reiher, M. (2015). An efficient matrix product operator representation of the quantum chemical hamiltonian. *J. Chem. Phys.* 143: 244118.

82 Chan, G.K.-L. (2008). Density matrix renormalization group lagrangians. *Phys. Chem. Chem. Phys.* 10: 3454–3459.

83 White, S.R. (2005). Density matrix renormalization group algorithms with a single center site. *Phys. Rev. B* 72: 180403.

84 Dorando, J.J., Hachmann, J., and Chan, G.K.-L. (2007). Targeted excited state algorithms. *J. Chem. Phys.* 127: 084109.

85 Ren, J., Yi, Y., and Shuai, Z. (2016). Inner space perturbation theory in matrix product states: replacing expensive iterative diagonalization. *J. Chem. Theory Comput.* 12: 4871–4878.

86 Guo, S., Li, Z., and Chan, G.K.-L. (2018). A perturbative density matrix renormalization group algorithm for large active spaces. *J. Chem. Theory Comput.* 14: 4063–4071.

87 Legeza, O., Noack, R.M., Sólyom, J., and Tincani, L. (2008). *Computational Many-Particle Physics*, Lecture Notes in Physics 739 (eds. H. Fehske, R. Schneider and A. Weiße), 653–664. Springer Berlin Heidelberg.

88 Legeza, O. and Sólyom, J. (2003). Optimizing the density-matrix renormalization group method using quantum information entropy. *Phys. Rev. B* 68: 195116.

89 Rissler, J., Noack, R.M., and White, S.R. (2006). Measuring orbital interaction using quantum information theory. *Chem. Phys.* 323: 519–531.

90 Boguslawski, K., Tecmer, P., Legeza, O., and Reiher, M. (2012). Entanglement measures for single- and multireference correlation effects. *J. Phys. Chem. Lett.* 3: 3129–3135.

91 Boguslawski, K., Tecmer, P., Barcza, G. et al. (2013). Orbital entanglement in bond-formation processes. *J. Chem. Theory Comput.* 9: 2959–2973.

92 Boguslawski, K. and Tecmer, P. (2015). Orbital entanglement in quantum chemistry. *Int. J. Quantum Chem.* 115: 1289–1295.

93 Boguslawski, K. and Tecmer, P. (2017). Erratum: orbital entanglement in quantum chemistry. *Int. J. Quantum Chem.* 117: e25455.

94 Legeza, O., Noack, R.M., Sólyom, J., and Tincani, L. (2008). *Computational Many-Particle Physics*, Lecture Notes in Physics 739 (eds. H. Fehske, R. Schneider and A. Weiße), 653–664. Springer Berlin Heidelberg.

95 Freitag, L., Knecht, S., Keller, S.F. et al. (2015). Orbital entanglement and CASSCF analysis of the Ru–NO bond in a ruthenium nitrosyl complex. *Phys. Chem. Chem. Phys.* 17: 14383–14392.

96 Stein, C.J. and Reiher, M. (2016). Automated selection of active orbital spaces. *J. Chem. Theory Comput.* 12: 1760–1771.

97 Stein, C.J., von Burg, V., and Reiher, M. (2016). The delicate balance of static and dynamic electron correlation. *J. Chem. Theory Comput.* 12: 3764–3773.

98 Stein, C.J. and Reiher, M. (2019). autoCAS: a program for fully automated multi-configurational calculations. *J. Comput. Chem.* 40: 2216–2226.

99 Sinha, S.B., Shopov, D.Y., Sharninghausen, L.S. et al. (2017). Redox activity of oxo-bridged iridium dimers in an N,O-donor environment: characterization of remarkably stable Ir(IV,V) complexes. *J. Am. Chem. Soc.* 139: 9672–9683.

100 Becker, P., Duhamel, T., Stein, C.J. et al. (2017). Cooperative light-activated iodine and photoredox catalysis for the amination of C–H bonds. *Angew. Chem. Int. Ed.* 56: 8004–8008.

101 Husch, T., Freitag, L., and Reiher, M. (2018). Calculation of ligand dissociation energies in large transition-metal complexes. *J. Chem. Theory Comput.* 14: 2456–2468.

102 Stein, C.J. and Reiher, M. (2017). Automated identification of relevant frontier orbitals for chemical compounds and processes. *Chimia* 71: 170–176.

103 Stein, C.J. and Reiher, M. (2017). Measuring multi-configurational character by orbital entanglement. *Mol. Phys.* 115: 2110–2119.

104 Nakatani, N., Wouters, S., Van Neck, D., and Chan, G.K.-L. (2014). Linear response theory for the density matrix renormalization group: efficient algorithms for strongly correlated excited states. *J. Chem. Phys.* 140: 024108.

105 McCulloch, I.P. (2007). From density-matrix renormalization group to matrix product states. *J. Stat. Mech. Theory Exp.* 2007: P10014.

106 Keller, S.F. and Reiher, M. (2014). Determining factors for the accuracy of DMRG in chemistry. *Chimia* 68: 200–203.

107 Moritz, G. and Reiher, M. (2007). Decomposition of density matrix renormalization group states into a Slater determinant basis. *J. Chem. Phys.* 126: 244109.

108 Barcza, G., Legeza, O., Marti, K.H., and Reiher, M. (2011). Quantum-information analysis of electronic states of different molecular structures. *Phys. Rev. A* 83: 012508.

109 Ghosh, D., Hachmann, J., Yanai, T., and Chan, G.K.-L. (2008). Orbital optimization in the density matrix renormalization group, with applications to polyenes and β-carotene. *J. Chem. Phys.* 128: 144117.

110 Ma, Y., Knecht, S., Keller, S., and Reiher, M. (2017). Second-order self-consistent-field density-matrix renormalization group. *J. Chem. Theory Comput.* 13: 2533–2549.

111 Sun, Q., Yang, J., and Chan, G.K.-L. (2017). A general second order complete active space self-consistent-field solver for large-scale systems. *Chem. Phys. Lett.* 683: 291–299.

112 Legeza, O. and Sólyom, J. (2004). Quantum data compression, quantum information generation, and the density-matrix renormalization-group method. *Phys. Rev. B* 70: 205118.

113 Yanai, T., Kurashige, Y., Mizukami, W. et al. (2015). Density matrix renormalization group for ab initio calculations and associated dynamic correlation methods: a review of theory and applications. *Int. J. Quantum Chem.* 115: 283–299.

114 Kurashige, Y. and Yanai, T. (2011). Second-order perturbation theory with a density matrix renormalization group self-consistent field reference function: theory and application to the study of chromium dimer. *J. Chem. Phys.* 135: 094104.

115 Kurashige, Y. (2014). Multireference electron correlation methods with density matrix renormalisation group reference functions. *Mol. Phys.* 112: 1485–1494.

116 Nakatani, N. and Guo, S. (2017). Density matrix renormalization group (DMRG) method as a common tool for large active-space CASSCF/CASPT2 calculations. *J. Chem. Phys.* 146: 094102.

117 Yanai, T., Saitow, M., Xiong, X.-G. et al. (2017). Multistate complete-active-space second-order perturbation theory based on density matrix renormalization group reference states. *J. Chem. Theory Comput.* 13: 4829–4840.

118 Guo, S., Watson, M.A., Hu, W. et al. (2016). N-electron valence state perturbation theory based on a density matrix renormalization group reference function, with applications to the chromium dimer and a trimer model of poly(p-phenylenevinylene). *J. Chem. Theory Comput.* 12: 1583–1591.

119 Freitag, L., Knecht, S., Angeli, C., and Reiher, M. (2017). Multireference perturbation theory with Cholesky decomposition for the density matrix renormalization group. *J. Chem. Theory Comput.* 13: 451–459.

120 Mazziotti, D.A. (1998). Approximate solution for electron correlation through the use of schwinger probes. *Chem. Phys. Lett.* 289: 419–427.

121 Harris, F.E. (2002). Cumulant-based approximations to reduced density matrices. *Int. J. Quantum Chem.* 90: 105–113.

122 Kutzelnigg, W., Shamasundar, K.R., and Mukherjee, D. (2010). Spinfree formulation of reduced density matrices, density cumulants and generalised normal ordering. *Mol. Phys.* 108: 433–451.

123 Zgid, D., Ghosh, D., Neuscamman, E., and Chan, G.K.-L. (2009). *A Study of Cumulant Approximations to N-Electron Valence Multireference Perturbation Theory*, vol. 130, 194107.

124 Kurashige, Y., Chalupský, J., Lan, T.N., and Yanai, T. (2014). Complete active space second-order perturbation theory with cumulant approximation for extended active-space wave function from density matrix renormalization group. *J. Chem. Phys.* 141: 174111.

125 Saitow, M., Kurashige, Y., and Yanai, T. (2013). Multireference configuration interaction theory using cumulant reconstruction with internal contraction of density matrix renormalization group wave function. *J. Chem. Phys.* 139: 044118.

126 Saitow, M., Kurashige, Y., and Yanai, T. (2015). Fully internally contracted multireference configuration interaction theory using density matrix renormalization group: a reduced-scaling implementation derived by computer-aided tensor factorization. *J. Chem. Theory Comput.* 11: 5120–5131.

127 Luo, Z., Ma, Y., Wang, X., and Ma, H. (2018). Externally-contracted multireference configuration interaction method using a DMRG reference wave function. *J. Chem. Theory Comput.* 14: 4747–4755.

128 Phung, Q.M., Wouters, S., and Pierloot, K. (2016). Cumulant approximated second-order perturbation theory based on the density matrix renormalization group for transition metal complexes: a benchmark study. *J. Chem. Theory Comput.* 12: 4352–4361.

129 Sharma, S., Jeanmairet, G., and Alavi, A. (2016). Quasi-degenerate perturbation theory using matrix product states. *J. Chem. Phys.* 144: 034103.

130 Sharma, S. and Chan, G.K.-L. (2014). Communication: a flexible multi-reference perturbation theory by minimizing the Hylleraas functional with matrix product states. *J. Chem. Phys.* 141: 111101.

131 Sharma, S., Knizia, G., Guo, S., and Alavi, A. (2017). Combining internally contracted states and matrix product states to perform multireference perturbation theory. *J. Chem. Theory Comput.* 13: 488–498.

132 Sokolov, A.Y., Guo, S., Ronca, E., and Chan, G.K.-L. (2017). Time-dependent n-electron valence perturbation theory with matrix product state reference wave functions for large active spaces and basis sets: applications to the chromium dimer and all-trans polyenes. *J. Chem. Phys.* 146: 244102.

133 Sokolov, A.Y. and Chan, G.K.-L. (2016). A time-dependent formulation of multi-reference perturbation theory. *J. Chem. Phys.* 144: 064102.

134 Roemelt, M., Guo, S., and Chan, G.K.-L. (2016). A projected approximation to strongly contracted n-electron valence perturbation theory for DMRG wave functions. *J. Chem. Phys.* 144: 204113.

135 Yanai, T. and Chan, G.K.-L. (2006). Canonical transformation theory for multireference problems. *J. Chem. Phys.* 124: 194106.

136 Yanai, T. and Chan, G.K.-L. (2007). Canonical transformation theory from extended normal ordering. *J. Chem. Phys.* 127: 104107.

137 Chan, G.K.-L. and Yanai, T. (2007). *Reduced-Density-Matrix Mechanics: With Application to Many-Electron Atoms and Molecules* (ed. D. Mazziotti), 343–384. Wiley.

138 Yanai, T., Kurashige, Y., Neuscamman, E., and Chan, G.K.-L. (2010). Multireference quantum chemistry through a joint density matrix renormalization group and canonical transformation theory. *J. Chem. Phys.* 132: 024105.

139 Neuscamman, E., Yanai, T., and Chan, G.K.-L. (2010). Strongly contracted canonical transformation theory. *J. Chem. Phys.* 132: 024106.

140 Neuscamman, E., Yanai, T., and Chan, G.K.-L. (2010). A review of canonical transformation theory. *Int. Rev. Phys. Chem.* 29: 231–271.

141 Yanai, T., Kurashige, Y., Neuscamman, E., and Chan, G.K.-L. (2012). Extended implementation of canonical transformation theory: parallelization and a new level-shifted condition. *Phys. Chem. Chem. Phys.* 14: 7809–7820.

142 Leininger, T., Stoll, H., Werner, H.-J., and Savin, A. (1997). Combining long-range configuration interaction with short-range density functionals. *Chem. Phys. Lett.* 275: 151–160.

143 Toulouse, J., Colonna, F., and Savin, A. (2004). Long-range–short-range separation of the electron-electron interaction in density-functional theory. *Phys. Rev. A* 70: 062505.

144 Fromager, E., Toulouse, J., and Jensen, H.J.A. (2007). On the universality of the long-/short-range separation in multiconfigurational density-functional theory. *J. Chem. Phys.* 126: 074111.

145 Fromager, E., Réal, F., Wåhlin, P. et al. (2009). On the universality of the long-/short-range separation in multiconfigurational density-functional theory. II. Investigating f^0 actinide species. *J. Chem. Phys.* 131: 054107.

146 Hedegård, E.D., Toulouse, J., and Jensen, H.J.A. (2018). Multiconfigurational short-range density-functional theory for open-shell systems. *J. Chem. Phys.* 148: 214103.

147 Hedegård, E.D., Knecht, S., Kielberg, J.S. et al. (2015). Density matrix renormalization group with efficient dynamical electron correlation through range separation. *J. Chem. Phys.* 142: 224108.

148 Veis, L., Antalík, A., Brabec, J. et al. (2016). Coupled cluster method with single and double excitations tailored by matrix product state wave functions. *J. Phys. Chem. Lett.* 7: 4072–4078.

149 Faulstich, F.M., Laestadius, A., Kvaal, S. et al. (2019). Analysis of the coupled-cluster method tailored by tensor-network states in quantum chemistry. *SIAM J. Numer. Anal.* 57: 2579–2607.

150 Sharma, S. and Alavi, A. (2015). Multireference linearized coupled cluster theory for strongly correlated systems using matrix product states. *J. Chem. Phys.* 143: 102815.

151 Dresselhaus, T., Neugebauer, J., Knecht, S. et al. (2015). Self-consistent embedding of density-matrix renormalization group wave functions in a density functional environment. *J. Chem. Phys.* 142: 044111.

152 Hedegård, E.D. and Reiher, M. (2016). Polarizable embedding density matrix renormalization group. *J. Chem. Theory Comput.* 12: 4242–4253.

153 Baerends, E.J., Ellis, D.E., and Ros, P. (1973). Self-consistent molecular Hartree–Fock–Slater calculations I. The computational procedure. *Chem. Phys.* 2: 41–51.

154 Whitten, J.L. (1973). Coulombic potential energy integrals and approximations. *J. Chem. Phys.* 58: 4496–4501.

155 Vahtras, O., Almlöf, J., and Feyereisen, M.W. (1993). Integral approximations for LCAO-SCF calculations. *Chem. Phys. Lett.* 213: 514–518.

156 Ten-no, S. and Iwata, S. (1996). Multiconfiguration self-consistent field procedure employing linear combination of atomic-electron distributions. *J. Chem. Phys.* 105: 3604–3611.

157 Ten-No, S. and Iwata, S. (1996). On approximating electron repulsion integrals with linear combination of atomic-electron distributions. *Int. J. Quantum Chem.* 60: 1319–1324.

158 Beebe, N.H.F. and Linderberg, J. (1977). Simplifications in the generation and transformation of two-electron integrals in molecular calculations. *Int. J. Quantum Chem.* 12: 683–705.

159 Aquilante, F., Pedersen, T.B., Lindh, R. et al. (2008). Accurate ab initio density fitting for multiconfigurational self-consistent field methods. *J. Chem. Phys.* 129: 024113.

160 Aquilante, F., Malmqvist, P.-A., Pedersen, T.B. et al. (2008). Cholesky decomposition-based multiconfiguration second-order perturbation theory (CD-CASPT2): application to the spin-state energetics of Co^{III} (diiminato)(NPh). *J. Chem. Theory Comput.* 4: 694–702.

161 Guo, Y., Sivalingam, K., Valeev, E.F., and Neese, F. (2016). SparseMaps—A systematic infrastructure for reduced-scaling electronic structure methods. III. Linear-scaling multireference domain-based pair natural orbital n-electron valence perturbation theory. *J. Chem. Phys.* 144: 094111.

162 Stålring, J., Bernhardsson, A., and Lindh, R. (2001). Analytical gradients of a state average mcscf state and a state average diagnostic. *Mol. Phys.* 99: 103–114.

163 Snyder, J.W., Hohenstein, E.G., Luehr, N., and Martínez, T.J. (2015). An atomic orbital-based formulation of analytical gradients and nonadiabatic coupling vector elements for the state-averaged complete active space self-consistent field method on graphical processing units. *J. Chem. Phys.* 143: 154107.

164 Snyder, J.W., Fales, B.S., Hohenstein, E.G. et al. (2017). A direct-compatible formulation of the coupled perturbed complete active space self-consistent field equations on graphical processing units. *J. Chem. Phys.* 146: 174113.

165 Dupuis, M. (1981). Energy derivatives for configuration interaction wave functions. *J. Chem. Phys.* 74: 5758–5765.

166 Osamura, Y., Yamaguchi, Y., and Schaefer, H.F. (1982). Generalization of analytic configuration interaction (ci) gradient techniques for potential energy hypersurfaces, including a solution to the coupled perturbed hartree–fock equations for multiconfiguration SCF molecular wave functions. *J. Chem. Phys.* 77: 383–390.

167 Hu, W. and Chan, G.K.-L. (2015). Excited-state geometry optimization with the density matrix renormalization group, as applied to polyenes. *J. Chem. Theory Comput.* 11: 3000–3009.

168 Freitag, L., Ma, Y., Baiardi, A. et al. (2019). Approximate analytical gradients and nonadiabatic couplings for the state-average density matrix renormalization group self-consistent-field method. *J. Chem. Theory Comput.* 15: 6724–6737.

169 Lengsfield, B.H. III,, Saxe, P., and Yarkony, D.R. (1984). On the evaluation of nonadiabatic coupling matrix elements using SA-MCSCF/CI wave functions and analytic gradient methods. *I. J. Chem. Phys.* 81: 4549–4553.

170 Lengsfield, B.H. and Yarkony, D.R. (2007). Nonadiabatic interactions between potential energy surfaces: theory and applications. *Adv. Chem. Phys.*: 1–71.

171 Fdez. Galván, I., Delcey, M.G., Pedersen, T.B. et al. (2016). Analytical state-average complete-active-space self-consistent field nonadiabatic coupling vectors: implementation with density-fitted two-electron integrals and application to conical intersections. *J. Chem. Theory Comput.* 12: 3636–3653.

172 Eisert, J. (2013). *Emergent Phenomena in Correlated Matter* (eds. E. Pavarini, E. Koch and U. Schollwöck), 17.1–17.41. Verlag: Forschungszentrum Jülich.

173 Orús, R. (2014). A practical introduction to tensor networks: matrix product states and projected entangled pair states. *Ann. Phys.* 349: 117–158.

174 Ran, S.-J., Tirrito, E., Peng, C., Chen, X., Su, G., Lewenstein, M. (2017) Review of Tensor Network Contraction Approaches, arXiv:1708.09213.

175 Bridgeman, J.C. and Chubb, C.T. (2017). Hand-waving and interpretive dance: an introductory course on tensor networks. *J. Phys. A Math. Theor.* 50: 223001.

176 Szalay, S., Pfeffer, M., Murg, V. et al. (2015). Tensor product methods and entanglement optimization for ab initio quantum chemistry. *Int. J. Quantum Chem.* 115: 1342–1391.

177 Marti, K.H., Bauer, B., Reiher, M. et al. (2010). Complete-graph tensor network states: a new fermionic wave function ansatz for molecules. *New J. Phys.* 12: 103008.

178 Kovyrshin, A. and Reiher, M. (2016). Tensor network states with three-site correlators. *New J. Phys.* 18: 113001.

179 Kovyrshin, A. and Reiher, M. (2017). Self-adaptive tensor network states with multi-site correlators. *J. Chem. Phys.* 147: 214111.

180 Murg, V., Verstraete, F., Legeza, O., and Noack, R.M. (2010). Simulating strongly correlated quantum systems with tree tensor networks. *Phys. Rev. B* 82: 205105.

181 Nakatani, N. and Chan, G.K.-L. (2013). Efficient tree tensor network states (TTNS) for quantum chemistry: generalizations of the density matrix renormalization group algorithm. *J. Chem. Phys.* 138: 134113.

182 Murg, V., Verstraete, F., Schneider, R. et al. (2015). Tree tensor network state with variable tensor order: an efficient multireference method for strongly correlated systems. *J. Chem. Theory Comput.* 11: 1027–1036.

183 Gunst, K., Verstraete, F., Wouters, S. et al. (2018). T3NS: Three-legged tree tensor network states. *J. Chem. Theory Comput.* 14: 2026–2033.

184 Hachmann, J., Dorando, J.J., Avilés, M., and Chan, G.K.-L. (2007). The radical character of the acenes: a density matrix renormalization group study. *J. Chem. Phys.* 127: 134309.

185 Mizukami, W., Kurashige, Y., and Yanai, T. (2013). More π electrons make a difference: emergence of many radicals on graphene nanoribbons studied by ab initio DMRG theory. *J. Chem. Theory Comput.* 9: 401–407.

186 Wouters, S., Limacher, P.A., Van Neck, D., and Ayers, P.W. (2012). Longitudinal static optical properties of hydrogen chains: finite field extrapolations of matrix product state calculations. *J. Chem. Phys.* 136: 134110.

187 Ma, Y. and Ma, H. (2013). Assessment of various natural orbitals as the basis of large active space density-matrix renormalization group calculations. *J. Chem. Phys.* 138: 224105.

188 Ronca, E., Li, Z., Jimenez-Hoyos, C.A., and Chan, G.K.-L. (2017). Time-step targeting time-dependent and dynamical density matrix renormalization group algorithms with ab initio hamiltonians. *J. Chem. Theory Comput.* 13: 5560–5571.

189 Ma, Y., Wen, J., and Ma, H. (2015). Density-matrix renormalization group algorithm with multi-level active space. *J. Chem. Phys.* 143: 034105.

190 Marti, K.H., Malkin Ondík, I., Moritz, G., and Reiher, M. (2008). Density matrix renormalization group calculations on relative energies of transition metal complexes and clusters. *J. Chem. Phys.* 128: 014104.

191 Yanai, T., Kurashige, Y., Ghosh, D., and Chan, G.K.-L. (2009). Accelerating convergence in iterative solution for large-scale complete active space self-consistent-field calculations. *Int. J. Quantum Chem.* 109: 2178–2190.

192 Mizukami, W., Kurashige, Y., and Yanai, T. (2010). Communication: novel quantum states of electron spins in polycarbenes from ab initio density matrix renormalization group calculations. *J. Chem. Phys.* 133: 091101.

193 Kawakami, T., Saito, T., Sharma, S. et al. (2017). Full-valence density matrix renormalisation group calculations on meta-benzyne based on unrestricted natural orbitals. Revisit of seamless continuation from broken-symmetry to symmetry-adapted models for diradicals. *Mol. Phys.* 115: 2267–2284.

194 Wouters, S., Van Speybroeck, V., and Van Neck, D. (2016). DMRG-CASPT2 study of the longitudinal static second hyperpolarizability of all-trans polyenes. *J. Chem. Phys.* 145: 054120.

195 Barford, W., Bursill, R.J., and Lavrentiev, M.Y. (2001). Density-matrix renormalization-group calculations of excited states of linear polyenes. *Phys. Rev. B* 63: 195108.

196 Ma, H., Cai, F., Liu, C., and Jiang, Y. (2005). Spin distribution in neutral polyene radicals: Pariser–Parr–Pople model studied with the density matrix renormalization group method. *J. Chem. Phys.* 122: 104909.

197 Ma, H., Liu, C., and Jiang, Y. (2005). Theoretical study of the lowest $\pi \rightarrow \pi^*$ excitation energies for neutral and doped polyenes. *J. Chem. Phys.* 123: 084303.

198 Bursill, R.J., Barford, W., and Daly, H. (1999). Molecular orbital models of benzene. Biphenyl and the oligophenylenes. *Chem. Phys.* 243: 35–44.

199 Raghu, C., Pati, Y.A., and Ramasesha, S. (2002). Structural and electronic instabilities in polyacenes: density-matrix renormalization group study of a long-range interacting model. *Phys. Rev. B* 65: 155204.

200 Das, M. and Ramasesha, S. (2006). A density matrix renormalization group study of low-lying excitations of polythiophene within a Pariser–Parr–Pople model. *J. Chem. Sci.* 118: 67–78.

201 Das, M. (2010). Low-lying excitations of poly-fused thiophene within Pariser–Parr–Pople model: a density matrix renormalization group study. *J. Chem. Phys.* 132: 194107.

202 Mukhopadhyay, S. and Ramasesha, S. (2009). Study of linear and nonlinear optical properties of dendrimers using density matrix renormalization group method. *J. Chem. Phys.* 131: 074111.

203 Goli, V.M.L.D.P., Prodhan, S., Mazumdar, S., and Ramasesha, S. (2016). Correlated electronic properties of some graphene nanoribbons: a DMRG study. *Phys. Rev. B* 94: 035139.

204 Kumar, M., Pati, Y.A., and Ramasesha, S. (2012). A density matrix renormalization group method study of optical properties of porphines and metalloporphines. *J. Chem. Phys.* 136: 014112.

205 Thomas, S., Pati, Y.A., and Ramasesha, S. (2013). Linear and nonlinear optical properties of expanded porphyrins: a DMRG study. *J. Phys. Chem. A* 117: 7804–7809.

206 Kurashige, Y., Chan, G.K.-L., and Yanai, T. (2013). Entangled quantum electronic wave functions of the Mn_4 CaO_5 cluster in photosystem II. *Nat. Chem.* 5: 660–666.

207 Kawakami, T., Miyagawa, K., Isobe, H. et al. (2018). Relative stability between the manganese hydroxide- and oxo-models for water oxidation by CCSD, DMRG CASCI, CASSCF, CASPT2 and CASDFT methods; Importance of static and dynamical electron correlation effects for OEC of PSII. *Chem. Phys. Lett.* 705: 85–91.

208 Roemelt, M., Krewald, V., and Pantazis, D.A. (2018). Exchange coupling interactions from the density matrix renormalization group and n-electron valence perturbation theory: application to a biomimetic mixed-valence manganese complex. *J. Chem. Theory Comput.* 14: 166–179.

209 Radoń, M. and Pierloot, K. (2008). Binding of CO, NO, and O_2 to heme by density functional and multireference ab initio calculations. *J. Phys. Chem. A* 112: 11824–11832.

210 Phung, Q.M. and Pierloot, K. (2018). The dioxygen adducts of iron and manganese porphyrins: electronic structure and binding energy. *Phys. Chem. Chem. Phys.* 20: 17009–17019.

211 Boguslawski, K., Marti, K.H., Legeza, O., and Reiher, M. (2012). Accurate ab initio spin densities. *J. Chem. Theory Comput.* 8: 1970–1982.

212 Chalupský, J., Rokob, T.A., Kurashige, Y. et al. (2014). Reactivity of the binuclear non-heme iron active site of Δ^9 desaturase studied by large-scale multireference ab initio calculations. *J. Am. Chem. Soc.* 136: 15977–15991.

213 Kurashige, Y., Saitow, M., Chalupský, J., and Yanai, T. (2014). Radical O–O coupling reaction in diferrate-mediated water oxidation studied using multireference wave function theory. *Phys. Chem. Chem. Phys.* 16: 11988–11999.

214 Zhao, Y., Boguslawski, K., Tecmer, P. et al. (2015). Dissecting the bond-formation process of d^{10}-metal–ethene complexes with multireference approaches. *Theor. Chem. Acc.* 134: 1–10.

215 Dong, G., Phung, Q.M., Hallaert, S.D. et al. (2017). H_2 binding to the active site of [NiFe] hydrogenase studied by multiconfigurational and coupled-cluster methods. *Phys. Chem. Chem. Phys.* 19: 10590–10601.

216 Artiukhin, D.G., Stein, C.J., Reiher, M., and Neugebauer, J. (2017). Quantum chemical spin densities for radical cations of photosynthetic pigment models. *Photochem. Photobiol.* 93: 815–833.

217 Harris, T.V., Kurashige, Y., Yanai, T., and Morokuma, K. (2014). Ab initio density matrix renormalization group study of magnetic coupling in dinuclear iron and chromium complexes. *J. Chem. Phys.* 140: 054303.

218 Nguyen Lan, T., Kurashige, Y., and Yanai, T. (2015). Scalar relativistic calculations of hyperfine coupling constants using ab initio density matrix renormalization group method in combination with third-order Douglas–Kroll–Hess transformation: case studies on 4*d* transition metals. *J. Chem. Theory Comput.* 11: 73–81.

219 Knecht, S., Keller, S., Autschbach, J., and Reiher, M. (2016). A nonorthogonal state-interaction approach for matrix product state wave functions. *J. Chem. Theory Comput.* 12: 5881–5894.

220 Battaglia, S., Keller, S., and Knecht, S. (2018). Efficient relativistic density-matrix renormalization group implementation in a matrix-product formulation. *J. Chem. Theory Comput.* 14: 2353–2369.

221 Sayfutyarova, E.R. and Chan, G.K.-L. (2018). Electron paramagnetic resonance g-tensors from state interaction spin-orbit coupling density matrix renormalization group. *J. Chem. Phys.* 148: 184103.

222 Roemelt, M. and Pantazis, D.A. (2019). Multireference approaches to spin-state energetics of transition metal complexes utilizing the density matrix renormalization group. *Adv. Theory Simul.*: 1800201.

223 König, E. and Madeja, K. (1967). Infra-red spectra at the $^5T_2$–$^1A_1$ cross-over in iron (II) complexes. *Spectrochim. Acta A Mol. Spectrosc.* 23: 45–54.

224 Reiher, M. (2002). Theoretical study of the $Fe(Phen)_2$ $(NCS)_2$ spin-crossover complex with reparametrized density functionals. *Inorg. Chem.* 41: 6928–6935.

225 Swart, M., Güell, M., and Solà, M. (2010). *Quantum Biochemistry* (ed. C. F. Matta), 551–583. Wiley-VCH Verlag GmbH & Co. KGaA.

226 Swart, M. (2013). Spin states of (bio)inorganic systems: successes and pitfalls. *Int. J. Quantum Chem.* 113: 2–7.

227 Costas, M. and Harvey, J.N. (2013). Spin states: discussion of an open problem. *Nat. Chem.* 5: 7–9.

228 Schröder, D., Shaik, S., and Schwarz, H. (2000). Two-state reactivity as a new concept in organometallic chemistry. *Acc. Chem. Res.* 33: 139–145.

229 Usharani, D., Wang, B., Sharon, D.A., and Shaik, S. (2015). *Spin States in Biochemistry and Inorganic Chemistry* (eds. M. Swart and M. Costas), 131–156. Wiley.

230 Wouters, S., Bogaerts, T., Voort, P.V.D. et al. (2014). Communication: DMRG-SCF study of the singlet, triplet, and quintet states of oxo-mn(salen). *J. Chem. Phys.* 140: 241103.

231 Phung, Q.M., Domingo, A., and Pierloot, K. (2017). Dinuclear iron(II) spin-crossover compounds: a theoretical study. *Chem. Eur. J.* 24: 5183–5190.

232 Coe, J.P., Almeida, N.M.S., and Paterson, M.J. (2017). Investigation of challenging spin systems using monte carlo configuration interaction and the density matrix renormalization group. *J. Comput. Chem.* 38: 2701–2712.

233 Pierloot, K., Phung, Q.M., and Domingo, A. (2017). Spin state energetics in first-row transition metal complexes: contribution of (3*s*3*p*) correlation and its description by second-order perturbation theory. *J. Chem. Theory Comput.* 13: 537–553.

234 Franz, K.J., Doerrer, L.H., Spingler, B., and Lippard, S.J. (2001). Pentacoordinate cobalt(III) thiolate and nitrosyl tropocoronand compounds. *Inorg. Chem.* 40: 3774–3780.

235 Hopmann, K.H., Conradie, J., Tangen, E. et al. (2015). Singlet–triplet gaps of cobalt nitrosyls: insights from tropocoronand complexes. *Inorg. Chem.* 54: 7362–7367.

236 Angeli, C., Cimiraglia, R., Evangelisti, S. et al. (2001). Introduction of n-electron valence states for multireference perturbation theory. *J. Chem. Phys.* 114: 10252.

237 Angeli, C., Cimiraglia, R., and Malrieu, J.-P. (2001). N-electron valence state perturbation theory: a fast implementation of the strongly contracted variant. *Chem. Phys. Lett.* 350: 297–305.

238 Angeli, C., Cimiraglia, R., and Malrieu, J.-P. (2002). N-electron valence state perturbation theory: a spinless formulation and an efficient implementation of the strongly contracted and of the partially contracted variants. *J. Chem. Phys.* 117: 9138–9153.

239 Knecht, S., Legeza, O., and Reiher, M. (2014). Communication: four-component density matrix renormalization group. *J. Chem. Phys.* 140: 041101.

240 Roemelt, M. (2015). Spin orbit coupling for molecular ab initio density matrix renormalization group calculations: application to g-tensors. *J. Chem. Phys.* 143: 044112.

241 Sayfutyarova, E.R. and Chan, G.K.-L. (2016). A state interaction spin-orbit coupling density matrix renormalization group method. *J. Chem. Phys.* 144: 234301.

242 King, H.F., Stanton, R.E., Kim, H. et al. (1967). Corresponding orbitals and the nonorthogonality problem in molecular quantum mechanics. *J. Chem. Phys.* 47: 1936–1941.

243 Malmqvist, P.-A. and Roos, B.O. (1989). The CASSCF state interaction method. *Chem. Phys. Lett.* 155: 189–194.

244 Boguslawski, K., Réal, F., Tecmer, P. et al. (2017). On the multi-reference nature of plutonium oxides:PuO_2^{2+}, PuO_2, PuO_3 and PuO_2 $(OH)_2$. *Phys. Chem. Chem. Phys.* 19: 4317–4329.

245 Tecmer, P., Boguslawski, K., Legeza, O., and Reiher, M. (2013). Unravelling the quantum-entanglement effect of noble gas coordination on the spin ground state of CUO. *Phys. Chem. Chem. Phys.* 16: 719–727.

246 Wouters, S., Poelmans, W., Ayers, P.W., and Van Neck, D. (2014). CheMPS2: a free open-source spin-adapted implementation of the density matrix renormalization group for ab initio quantum chemistry. *Comput. Phys. Commun.* 185: 1501–1514.

247 Liu, F., Kurashige, Y., Yanai, T., and Morokuma, K. (2013). Multireference ab initio density matrix renormalization group (DMRG)-CASSCF and DMRG-CASPT2 study on the photochromic ring opening of spiropyran. *J. Chem. Theory Comput.* 9: 4462–4469.

248 Ren, J., Peng, Q., Zhang, X. et al. (2017). Role of the dark ^2A_g state in donor–acceptor copolymers as a pathway for singlet fission: a DMRG study. *J. Phys. Chem. Lett.* 8: 2175–2181.

249 Pusch, S., Tröster, A., Lefrancois, D. et al. (2018). Mechanism and Cis/trans selectivity of vinylogous nazarov-type [6π] photocyclizations. *J. Org. Chem.* 83: 964–972.

250 Taffet, E.J., Olivier, Y., Lam, F. et al. (2018). Carbene–metal–amide bond deformation, rather than ligand rotation, drives delayed fluorescence. *J. Phys. Chem. Lett.* 9: 1620–1626.

251 Shirai, S., Kurashige, Y., and Yanai, T. (2016). Computational evidence of inversion of ^1L_a and ^1L_b-derived excited states in naphthalene excimer formation from ab initio multireference theory with large active space: DMRG-CASPT2 study. *J. Chem. Theory Comput.* 12: 2366–2372.

252 Ma, Y., Knecht, S., and Reiher, M. (2017). Multiconfigurational effects in theoretical resonance Raman spectra. *ChemPhysChem* 18: 384–393.

253 Heller, E.J., Sundberg, R., and Tannor, D. (1982). Simple aspects of Raman scattering. *J. Phys. Chem.* 86: 1822–1833.

254 Neugebauer, J. and Hess, B.A. (2004). Resonance Raman spectra of uracil based on Kramers–Kronig relations using time-dependent density functional calculations and multireference perturbation theory. *J. Chem. Phys.* 120: 11564–11577.

255 Mai, S., Richter, M., Marquetand, P., and González, L. (2014). *Excitation of Nucleobases from a Computational Perspective II: Dynamics; Topics in Current Chemistry*, 1–55. Springer Berlin Heidelberg.

256 Hassing, S. and Mortensen, O.S. (1980). Kramers–Kronig relations and resonance Raman scattering. *J. Chem. Phys.* 73: 1078–1083.

257 Peticolas, W.L. and Rush, T. (1995). Ab initio calculations of the ultraviolet resonance Raman spectra of uracil. *J. Comput. Chem.* 16: 1261–1270.

258 Rush, T.I. and Peticolas, W.L. (1995). Ab initio transform calculation of resonance Raman spectra of uracil, 1-methyluracil, and 5-methyluracil. *J. Phys. Chem.* 99: 14647–14658.

259 Sun, S. and Brown, A. (2014). Simulation of the resonance Raman spectrum for uracil. *J. Phys. Chem. A* 118: 9228–9238.

260 Mercier, Y., Santoro, F., Reguero, M., and Improta, R. (2008). The decay from the dark $n\pi^*$ excited state in uracil: an integrated CASPT2/CASSCF and PCM/TD-DFT study in the gas phase and in water. *J. Phys. Chem. B* 112: 10769–10772.

261 Baiardi, A., Stein, C.J., Barone, V., and Reiher, M. (2017). Vibrational density matrix renormalization group. *J. Chem. Theory Comput.* 13: 3764–3777.

262 Watson, J.K.G. (1968). Simplification of the molecular vibration-rotation hamiltonian. *Mol. Phys.* 15: 479–490.

263 Rakhuba, M. and Oseledets, I. (2016). Calculating vibrational spectra of molecules using tensor train decomposition. *J. Chem. Phys.* 145: 124101.

264 Knecht, S., Hedegård, E.D., Keller, S. et al. (2016). New approaches for ab initio calculations of molecules with strong electron correlation. *Chimia* 70: 244–251.
265 https://github.com/sanshar (accessed 14. 8. 2018).
266 https://github.com/SebWouters/CheMPS2 (accessed 14. 8. 2018).
267 https://scine.ethz.ch/download/ (accessed 14. 8. 2018).
268 https://github.com/naokin/mpsxx (accessed 14. 8. 2018).

8

Excited-State Calculations with Quantum Monte Carlo

Jonas Feldt and Claudia Filippi

MESA+ Institute for Nanotechnology, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands

Abstract

Quantum Monte Carlo methods are first-principle approaches that approximately solve the Schrödinger equation stochastically. As compared to traditional quantum chemistry methods, they offer important advantages such as the ability to handle a large variety of many-body wave functions, the favorable scaling with the number of particles, and the intrinsic parallelism of the algorithms which are particularly suitable to modern massively parallel computers. In this chapter, we focus on the two quantum Monte Carlo approaches most widely used for electronic structure problems, namely, the variational and diffusion Monte Carlo methods. We give particular attention to the recent progress in the techniques for the optimization of the wave function, a challenging and important step to achieve accurate results in both the ground and the excited state. We conclude with an overview of the current status of excited-state calculations for molecular systems, demonstrating the potential of quantum Monte Carlo methods in this field of applications.

8.1 Introduction

Quantum Monte Carlo (QMC) methods are a broad range of approaches which employ stochastic algorithms to simulate quantum systems. They have been used to study fermions and bosons at zero and finite temperature with very different many-body Hamiltonians and wave functions in the fields of molecular chemistry, condensed matter, and nuclear physics. While all QMC methods, despite the diversity of applications, share some core algorithms, we restrict ourselves here to the two zero-temperature continuum QMC methods[1] that are most commonly used in electronic structure theory, namely, variational (VMC) and diffusion (DMC) Monte Carlo [1–3].

As compared to deterministic quantum chemistry approaches, solving the Schrödinger equation by stochastic means in VMC or DMC offers several advantages. The stochastic nature of the integration allows for greater flexibility in the functional form of the many-body wave function employed, which can for instance include the explicit dependence on the inter-electronic distances. As a consequence, more compact wave functions can be used (the number of determinants needed to get the same energy is reduced by a few orders of magnitude) and, further, the dependence on the basis

1 A quantum Monte Carlo approach not in real space but in Slater determinant space (i.e., the full configuration interaction QMC method) is briefly introduced in Chapter 6.

Quantum Chemistry and Dynamics of Excited States: Methods and Applications,
First Edition. Edited by Leticia González and Roland Lindh.

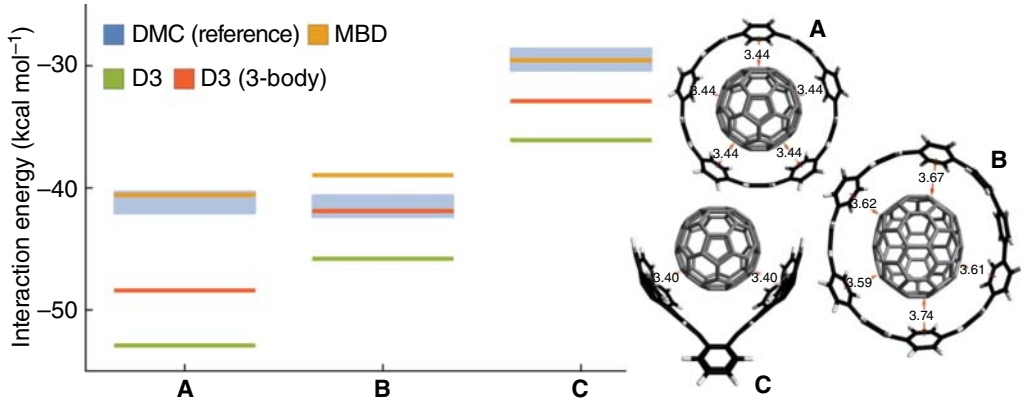


Figure 8.1 Interaction energies of the complexes **A**–**C** computed with DMC and density functional theory with two different dispersion corrections (D3) and the many-body dispersion method (MBD). Adapted from Ref. [4].

set is much weaker. Even though VMC and DMC are expensive, they have a favorable scaling with the system size (a mere polynomial N^4 in the number of electrons N), which has enabled simulations with hundreds and even thousands of electrons also in condensed matter, where traditional highly-correlated approaches are very difficult to apply. Finally, the intrinsically parallel nature of QMC algorithms renders these methods ideal candidates to take advantage of the massively parallel computers which are now available. An impressive example of such calculations is shown in Figure 8.1 where the interaction energy dominated by dispersion is benchmarked with DMC for remarkably large complexes [4].

That said, when inspecting the literature, it is evident that VMC and DMC methods have traditionally been employed to calculate mainly total energies and total energy differences as the computation of quantities other than the energy is more complicated. QMC calculations are, for instance, generally carried out on geometries obtained at a different level of theory and also the construction of the many-body wave function and its optimization are not straightforward. Particular care must in fact be paid to this step since the residual DMC error can be larger than sometimes assumed in the past, when calculations were anyhow limited to relatively simple wave functions and it was not feasible to extensively explore the dependence of the results on the choice of wave function. We will come back to this point later, which is especially relevant for excited states.

The last few years have, however, seen remarkable progress in methodological developments to overcome these and other limitations, as well as extend the applicability of QMC to larger systems both in the ground and the excited state. In particular, robust optimization algorithms for the parameters in the wave function have been developed for ground states [5–8] and extended to the state-average [9] as well as state-specific [10, 11] optimization of excited states. Importantly, it has recently become possible to efficiently compute the quantities needed in these optimization schemes (i.e., the derivatives of the wave function and the action of the Hamiltonian on these derivatives) at a cost per Monte Carlo step which scales like the computation of the energy alone [12–14]. Consequently, the determinantal component of a QMC wave function does not have to be borrowed from other quantum chemical calculations but can be consistently optimized within VMC after the addition of the correlation terms depending on the inter-electronic distances. These developments also enable the concomitant optimization of the structural parameters in VMC even when large determinantal expansions are employed in the wave function [14, 15]. The

possibility of performing molecular dynamics simulations with VMC forces has also been recently demonstrated [16, 17].

Researchers have also been actively investigating more complex functional forms [18–22] to recover missing correlation and allow a more compact wave function than the one obtained with a multi-determinant component. A local correlation description has also been shown to be a promising route to achieve smaller expansions and reduced computational costs for ground and excited states [23, 24]. In parallel, algorithms have been explored for a more automatic selection of the determinantal component, avoiding the possible pitfalls of a manual choice based on chemical intuition [15, 25–30]. Importantly, effort has been devoted to develop algorithms for the computation of quantities other than the energy via estimators characterized by reduced fluctuations as well as wave function bias [5, 31–36]. In addition to these methodological advances, various tools have become available to facilitate the calculations, such as tables of pseudopotentials and corresponding basis sets especially constructed for QMC [37–42]. Finally, multi-scale methods have been proposed to include the effects of a (responsive) environment on an embedded system treated with QMC [43–48].

After a brief description of the VMC and DMC methods, we will focus here on some of these recent developments, paying special attention to the algorithms employed to optimize the variational parameters in the wave function. We will then review relevant work and recent advances in the calculation of excited states and their properties, mainly for molecular systems. We note that useful sources for QMC are the introductory book to Monte Carlo methods and their use in quantum chemistry [49], and the reviews on QMC methods and their applications to solids [1, 50–52] and to non-covalent interactions [53]. A detailed introduction to VMC and DMC can be found in Ref. [54]. Finite-temperature path integral Monte Carlo methods are covered in Ref. [55].

8.2 Variational Monte Carlo

Variational Monte Carlo is the simplest flavor of QMC methods and represents a generalization of classical Monte Carlo to compute the multidimensional integrals in the expectation values of quantum mechanical operators. The approach enables the use of any "computable" wave function without severe restrictions on its functional form. This must be contrasted with other traditional quantum chemical methods which express the wave function as products of single particle orbitals in order to perform the relevant integrals analytically.

To illustrate how to compute an expectation value stochastically, let us consider the variational energy E_v, namely, the expectation value of the Hamiltonian $\hat{H}$ on a given wave function Ψ, which we rewrite as

$$E_\mathrm{v} = \frac{\int \Psi(\mathbf{R})^* \hat{H} \Psi(\mathbf{R}) \mathrm{d}\mathbf{R}}{\int |\Psi(\mathbf{R})|^2 \mathrm{d}\mathbf{R}} = \frac{\int |\Psi(\mathbf{R})|^2 \frac{\hat{H}\Psi(\mathbf{R})}{\Psi(\mathbf{R})} \mathrm{d}\mathbf{R}}{\int |\Psi(\mathbf{R})|^2 \mathrm{d}\mathbf{R}} = \int \rho(\mathbf{R}) E_\mathrm{L}(\mathbf{R}) \mathrm{d}\mathbf{R}\,, \tag{8.1}$$

where we have introduced the probability distribution, $\rho(\mathbf{R})$, and the local energy, $E_\mathrm{L}(\mathbf{R})$, defined as

$$\rho(\mathbf{R}) = \frac{|\Psi(\mathbf{R})|^2}{\int |\Psi(\mathbf{R})|^2 \mathrm{d}\mathbf{R}} \quad \text{and} \quad E_\mathrm{L}(\mathbf{R}) = \frac{\hat{H}\Psi(\mathbf{R})}{\Psi(\mathbf{R})}\,, \tag{8.2}$$

with $\mathbf{R}$ denoting the $3N$ coordinates of the electrons. We note that we can interpret $\rho(\mathbf{R})$ as a probability distribution since it is always non-negative and integrates to one.

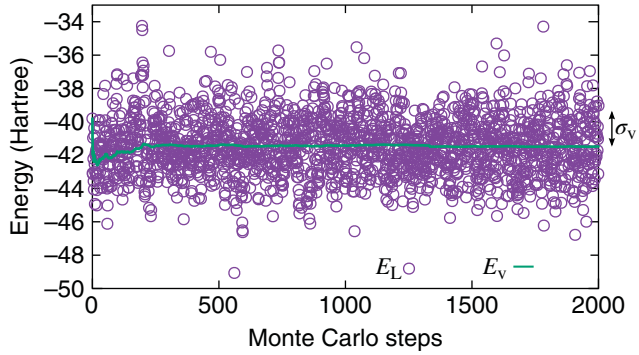


Figure 8.2 Local energy (circle) and its running average (green line) in a typical VMC run. The size of the root-mean-square fluctuations of the local energy, σ_v, is also indicated.

The integral can then be estimated by averaging the local energy computed on a set of M configurations $\{\mathbf{R}_k\}$ sampled from the probability density $\rho(\mathbf{R})$ as

$$E_v \approx \bar{E}_L = \frac{1}{M}\sum_k^M E_L(\mathbf{R}_k)\,. \tag{8.3}$$

According to the central limit theorem, this estimator converges to the exact value, E_v, with an increasing number of Monte Carlo configurations with a statistical uncertainty which decreases as

$$\mathrm{err}(\bar{E}_L) \propto \frac{\sigma_v}{\sqrt{M}}\,, \tag{8.4}$$

where $\sigma_v^2 = \int \rho(\mathbf{R})(E_L(\mathbf{R}) - E_v)^2 \mathrm{d}\mathbf{R}$ is the variance of the local energy. For this relation to hold, the chosen wave function must yield a finite variance of the sampled quantity, in this case, the local energy. A typical VMC run is illustrated in Figure 8.2, where the local energy is computed at each Monte Carlo step together with its running average.

Importantly, the statistical uncertainty decreases as $1/\sqrt{M}$ regardless the number of dimensions; therefore, Monte Carlo displays a faster convergence than deterministic numerical integration already for small numbers of dimensions.[2] Furthermore, as the trial wave function[3] improves, the Monte Carlo estimate of the variational energy requires fewer Monte Carlo steps to converge. In the limit of the wave function being an exact eigenstate of the Hamiltonian, the variance approaches zero and a single configuration is sufficient to obtain the exact variational energy. This *zero-variance* principle applies straightforwardly to the Hamiltonian and operators commuting with the Hamiltonian (therefore, the large number of total energy calculations found in the QMC literature). This principle can however be generalized to arbitrary observables by formulating an equivalent, improved estimator having the same average but a different, reduced variance [31]. Reduced variance estimators have been derived for the computation of electron density [56, 57], the electron pair densities [34], interatomic forces [33, 35], and other derivatives of the total energy [5].

In practice, the probability distribution $\rho(\mathbf{R})$ is sampled with the Metropolis–Hastings algorithm by simulating a Markov chain. This is a sequence of successive configurations, $\mathbf{R}_1, \ldots, \mathbf{R}_M$, generated with a transition probability, $P(\mathbf{R}'|\mathbf{R})$, where the transition to a new configuration $\mathbf{R}'$ only

2 For instance, the error for the Simpson's integration rule decreases as $1/M^{(4/d)}$ with d the number of dimensions and M the number of integration points, so Monte Carlo integration is already more efficient for $d > 8$.

3 A trial wave function is a wave function used as an approximation to the state of interest.

depends on the current point $\mathbf{R}$. The transition probability is stochastic, which means that it has the following properties:

$$P(\mathbf{R}'|\mathbf{R}) \geq 0 \quad \text{and} \quad \int P(\mathbf{R}'|\mathbf{R})\mathrm{d}\mathbf{R}' = 1\,. \tag{8.5}$$

Repeated application of P generates a Markov chain which converges to the target distribution ρ as

$$\lim_{M\to\infty} \int P(\mathbf{R}|\mathbf{R}_M)\ldots P(\mathbf{R}_2|\mathbf{R}_1)\rho_{\text{init}}(\mathbf{R}_1)\mathrm{d}\mathbf{R}_1\ldots\mathbf{R}_M = \rho(\mathbf{R})\,, \tag{8.6}$$

if P is ergodic (it is possible to move between two different configurations in a finite number of steps) and fulfills the so-called stationarity condition:

$$\int P(\mathbf{R}'|\mathbf{R})\rho(\mathbf{R})\mathrm{d}\mathbf{R} = \rho(\mathbf{R}')\,. \tag{8.7}$$

The stationarity condition tells us that, if we start from the desired distribution ρ, we will continue to sample ρ. Moreover, if the stochastic probability P is ergodic, it is possible to show that this condition ensures that any initial distribution will evolve to ρ under repeated applications of P.

In the Metropolis–Hastings algorithm, the transition to a new state is carried out in two steps: a new configuration is generated by a (stochastic) proposal probability and the proposed step is then accepted or rejected with an acceptance probability. The latter can be constructed so that the combined proposal and acceptance steps fulfill the stationarity condition. Most importantly, the acceptance depends only on ratios of $\rho(\mathbf{R})$ so that the generally unknown normalization of the distribution ρ is not required. We note that it is desirable to reduce sequential correlation among configurations. Proposing large steps to quickly explore the phase space must therefore be balanced against the rate of acceptance which decreases with large steps. For these reasons, electrons are generally moved one at the time to allow larger steps with a reasonable acceptance rate – this is a necessary feature as the system size grows since the size of the move would need to be decreased to have a reasonable acceptance of a move of all particles.

VMC is a very powerful method as the stochastic nature of integration gives a lot of freedom in the choice of the functional form of the wave function. It also allows us to learn a great deal about a system by exploring which ingredients in the wave function are necessary for its accurate description. Finally, in VMC, there is no sign problem associated with Fermi statistics, which generally plagues other quantum Monte Carlo approaches as we will see below. The obvious drawback is that, for each particular problem, a parametrization of the wave function has to be constructed. This process can be non-trivial and tends to be biased towards simpler electronic states: for example, it is easier to construct a good wave function for a closed-shell than an open-shell system so that the energy of the former will be closer to the exact result than the latter. Furthermore, properties other than the energy (or expectation values of operators commuting with the Hamiltonian) can be significantly less accurate since they are first order in the error of the wave function instead of second order as for the energy. It has, however, been shown that it is possible to extend this favorable property of the energy to arbitrary observables by using modified estimators which lead not only to reduced fluctuations but also to a reduced bias due to the wave function [33] as convincingly demonstrated in some promising applications [33, 34, 36, 57]. An example of this so-called zero-variance (ZV) zero-bias (ZB) approach applied to the computation of the intracule density is shown in Figure 8.3: the use of a ZV estimator significantly reduces the statistical fluctuations of the density and the further ZB formulation yields the correct result even when a simple Hartree–Fock wave function is employed. In general, the VMC approach is an extremely valuable tool and, in recent years, its use and impact has in fact become greater thanks to the availability of robust methods to optimize the many parameters in the wave function and, consequently, to increase the accuracy of the

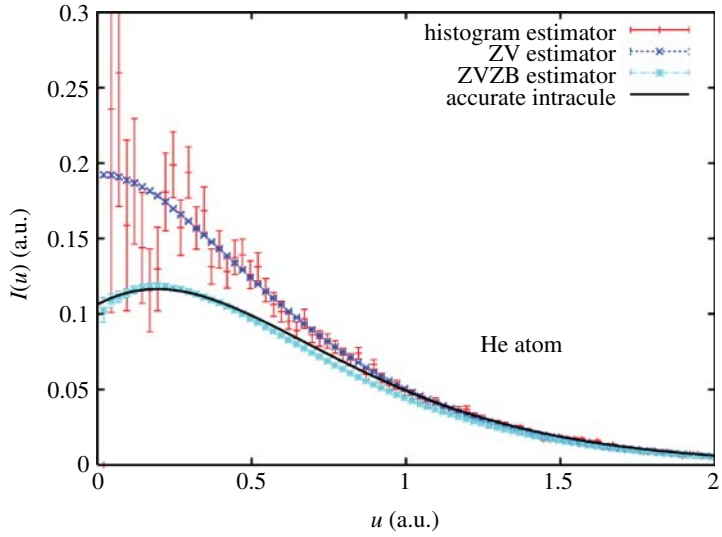


Figure 8.3 Spherically-averaged intracule density $I(u)$ as a function of the electron–electron distance u for the He atom calculated in VMC with a histogram, a zero-variance, and a zero-variance zero-bias estimator and the same Hartree–Fock wave function (without a Jastrow factor). Adapted from Ref. [34].

observables of interest already at the VMC level. Finally, characterizing and optimizing the trial wave function in VMC represents a necessary ingredient for more advanced projector Monte Carlo methods like the diffusion Monte Carlo approach described in the next Section.

8.3 Diffusion Monte Carlo

Projector Monte Carlo methods are QMC approaches which remove (at least in part) the bias of the trial wave function which characterizes VMC calculations. They are a stochastic implementation of the power method for finding the dominant eigenstate of a matrix or integral kernel. In a projector Monte Carlo method, one uses an operator that inverts the spectrum of $\hat{H}$ to project out the ground state of $\hat{H}$ from a given trial state.

Diffusion Monte Carlo (DMC) uses the exponential projection operator $e^{-t(\hat{H}-E_T)}$ with E_T a trial energy whose role will become immediately apparent. To understand the effect of applying this operator on a given wave function, let us consider a trial wave function Ψ, which we expand on the eigenstates of $\hat{H}$, Ψ_i with eigenvalues E_i. In the limit of infinite time t, we then obtain

$$\lim_{t\to\infty} e^{-t(\hat{H}-E_T)}|\Psi\rangle = \lim_{t\to\infty} \sum_i e^{-t(E_i-E_T)}|\Psi_i\rangle\langle\Psi_i|\Psi\rangle = \lim_{t\to\infty} e^{-t(E_0-E_T)}|\Psi_0\rangle\langle\Psi_0|\Psi\rangle\,, \tag{8.8}$$

where, in the last equality, we used that the coefficients in front of the higher eigenstates decay exponentially faster than that of the ground state. If we adjust E_T to E_0, the projection will yield the ground state Ψ_0. Note that the starting wave function must have a non-zero overlap with the ground-state one.

In the position representation, this projection can be rewritten as

$$\Psi(\mathbf{R}', t) = \int G(\mathbf{R}'|\mathbf{R}, t)\Psi(\mathbf{R})\mathrm{d}\mathbf{R}\,, \tag{8.9}$$

where we introduced the coordinate Green's function defined as

$$G(\mathbf{R}'|\mathbf{R}, t) = \langle\mathbf{R}'|e^{-t(\hat{H}-E_T)}|\mathbf{R}\rangle\,. \tag{8.10}$$

This representation readily allows us to see how to translate the projection into a Markov process provided that we can sample the Green's function and the trial wave function. For fermions, the wave function is antisymmetric and cannot therefore be interpreted as a probability distribution, a fact that we will ignore for the moment.

A further complication is that the exact form of the Green's function is not known. Fortunately, in the limit of small time steps τ, Trotter's theorem tells us that we are allowed to consider the potential and kinetic energy contributions separately since

$$e^{-\tau(\hat{T}+\hat{V})} = e^{-\tau\hat{T}}e^{-\tau\hat{V}} + \mathcal{O}(\tau^2)\,, \tag{8.11}$$

so that

$$\begin{aligned}\langle \mathbf{R}'|e^{-\tau\hat{H}}|\mathbf{R}\rangle \approx \langle \mathbf{R}'|e^{-\tau\hat{T}}e^{-\tau\hat{V}}|\mathbf{R}\rangle &= \int \mathrm{d}\mathbf{R}''\langle \mathbf{R}'|e^{-\tau\hat{T}}|\mathbf{R}''\rangle\langle \mathbf{R}''|e^{-\tau\hat{V}}|\mathbf{R}\rangle \\ &= \langle \mathbf{R}'|e^{-\tau\hat{T}}|\mathbf{R}\rangle e^{-\tau V(\mathbf{R})}\,. \end{aligned} \tag{8.12}$$

Therefore, we can rewrite the Green's function in the *short-time approximation* as

$$G(\mathbf{R}'|\mathbf{R},\tau) = (2\pi\tau)^{-3N/2}\ \exp\left[-\frac{(\mathbf{R}'-\mathbf{R})^2}{2\tau}\right]\ \exp[-\tau\ (V(\mathbf{R})-E_{\mathrm{T}})] + \mathcal{O}(\tau^2)\,, \tag{8.13}$$

where the first (stochastic) factor is the Green's function for diffusion while the second term multiplies the distribution by a positive scalar. The repeated application of the short-time Green's function to obtain the distribution at longer times (Eq. 8.9) can be interpreted as a Markov process with the difference that the Green's function is not normalized due to the potential term, and we therefore obtain a weighted random walk.

The basic DMC algorithm is rather simple:

1. An initial set of M_0 so-called walkers $\mathbf{R}_1, \ldots, \mathbf{R}_{M_0}$ is generated by sampling the trial wave function $\Psi(\mathbf{R})$ with the Metropolis algorithm as in VMC. This is the zeroth *generation* and the number of configurations is the *population* of the zeroth generation.
2. Each walker diffuses as $\mathbf{R}' = \mathbf{R} + \xi$ where ξ is sampled from the $3N$-dimensional Gaussian distribution $g(\xi) = (2\pi\tau)^{-3N/2}\exp(-\xi^2/2\tau)$.
3. For each walker, we compute the factor

$$p = \exp[-\tau(V(\mathbf{R}) - E_{\mathrm{T}})]\,. \tag{8.14}$$

   and perform the so-called branching step, namely, we branch the walker by treating p as the probability to survive at the next step: if $p < 1$, the walker survives with probability p while, if $p > 1$, the walker continues and new walkers with the same coordinates are created with probability $p - 1$. This is achieved by creating a number of copies of the current walker equal to the integer part of $p + \eta$ where η is a random number between (0,1). The branching step causes walkers to live in regions with a low potential $V < E_{\mathrm{T}}$ and die in regions with high V.
4. We adjust E_{T} so that the overall population fluctuates around the target value M_0.

Steps 2–4 are repeated until a stationary distribution is obtained and the desired properties are converged within a given statistical accuracy. A schematic representation of the evolution for a simple 1-dimensional problem is shown in Figure 8.4. Since the short-time expression of the Green's function is only valid in the limit of τ approaching zero, in practice, DMC calculations must be performed for different values of τ and the results extrapolated for τ which goes to zero.

The direct sampling of this Green's function proves, however, to be highly inefficient and unstable since the potential can vary significantly from configuration to configuration or also be unbounded like the Coulomb potential. For example, the electron-nucleus potential diverges to minus infinity as the two particles approach each other, and the branching factor will give rise to an unlimited number of walkers. Even if the potential is bounded, the approach becomes inefficient with

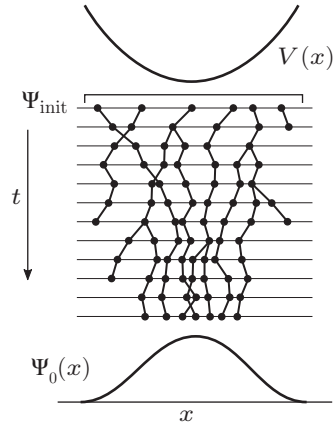


Figure 8.4 Schematic representation of a DMC simulation showing the evolution of the walkers in a 1-dimensional potential $V(x)$. The walkers are uniformly distributed at the start and converge to the ground-state distribution Ψ_0 after a number of time steps (adapted from Ref. [1]).

increasing size of the system since the branching factor also grows with the number of particles. These difficulties can be overcome by using *importance sampling*, where the trial wave function, Ψ, is used to guide the random walk. Starting from Eq. (8.9), we multiply each side by $\Psi(\mathbf{R}')$ and define the probability distribution $f(\mathbf{R}, t) = \Psi(\mathbf{R}, t)\Psi(\mathbf{R})$ which satisfies

$$f(\mathbf{R}', t) = \int \tilde{G}(\mathbf{R}'|\mathbf{R}, t)\Psi(\mathbf{R})^2 \mathrm{d}\mathbf{R}\,, \tag{8.15}$$

where the importance sampled Green's function is given by

$$\tilde{G}(\mathbf{R}'|\mathbf{R}, t) = \Psi(\mathbf{R}')\langle \mathbf{R}'|e^{-t(\hat{H}-E_{\mathrm{T}})}|\mathbf{R}\rangle/\Psi(\mathbf{R})\,. \tag{8.16}$$

In the limit of long times, this distribution $f(\mathbf{R}, t)$ approaches $\Psi_0(\mathbf{R})\Psi(\mathbf{R})$.

Assuming for the moment that $\Psi(\mathbf{R}')/\Psi(\mathbf{R}) > 0$, the importance sampled Green's function in the short-time approximation becomes

$$\tilde{G}(\mathbf{R}'|\mathbf{R}, \tau) \approx (2\pi\tau)^{-\frac{3}{2}N} \exp\left[-\frac{(\mathbf{R}' - \mathbf{R} - \tau\mathbf{V}(\mathbf{R}))^2}{2\tau}\right] \exp[-\tau(E_{\mathrm{L}}(\mathbf{R}) - E_{\mathrm{T}})]\,, \tag{8.17}$$

where one has assumed that the drift-velocity $\mathbf{V}(\mathbf{R}) = \nabla\Psi(\mathbf{R})/\Psi(\mathbf{R})$ and the local energy (Eq. 8.2) are constant in the step from $\mathbf{R}$ to $\mathbf{R}'$. There are two important new features of $\tilde{G}$. First, the quantum velocity $\mathbf{V}(\mathbf{R})$ pushes the walkers to regions where $\Psi(\mathbf{R})$ is large. In addition, the local energy E_{L} instead of the potential appears in the branching factor. Since the local energy becomes constant and equal to the eigenvalue as the trial wave function approaches the exact eigenstate, we expect that, for a good trial wave function, the fluctuations in the branching factor will be significantly smaller. In particular, imposing the cusp conditions on the wave function will remove the instabilities coming from the singular Coulomb potential. The simple DMC algorithm can be easily modified by sampling the square of the trial wave function in a VMC calculation (step 1), drifting before diffusing the walkers (step 2), and employing the exponential of the local energy as branching factor (step 3). Several important modifications to this bare-bones algorithm can and should be introduced to reduce the time-step error, which are described in detail along with further improvements in Ref. [58].

Up to this point, we have assumed that the wave function does not change sign. This is true for the ground state of a bosonic system, whose wave function can be in principle projected exactly in a DMC simulation. For fermions, however, a move of a walker can lead to a change of sign due to the

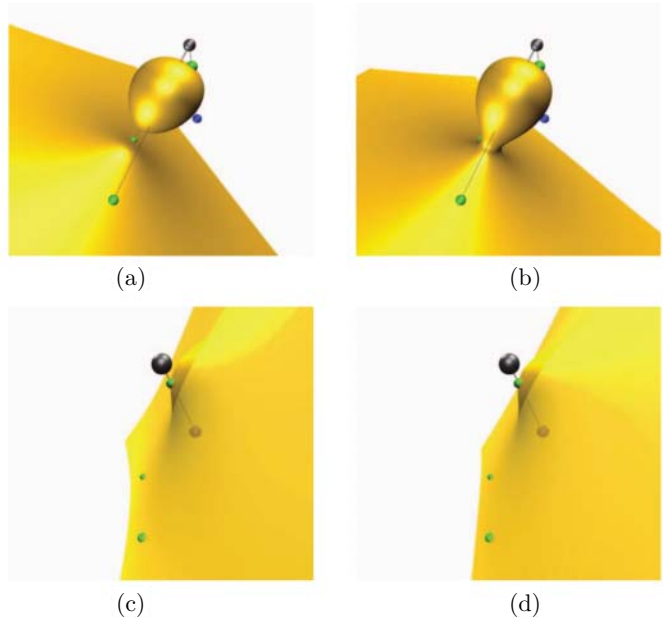


Figure 8.5 A 3-dimensional slice of the nodal surface of the N (top) and P (bottom) atoms obtained as a scan of the wave function moving one electron and keeping the others at snapshot positions (green/blue spheres). The black spheres mark the positions of the nucleus. (a, c) Nearly exact nodes and (b, d) Hartree–Fock nodes. Adapted from Ref. [59].

antisymmetry of the wave function. While it is possible to work with weights that carry a sign, the stochastic realization of such a straightforward approach is not stable since the separate evolution of the populations of positive and negative walkers will lead to the same bosonic solution, and the fermionic signal will be exponentially lost in the noise. This is known as the *fermionic sign problem*. To circumvent this problem, we can simply forbid moves in which the sign of the trial wave function changes and the walker crosses the nodes which are defined as $(3N-1)$-dimensional surfaces where the trial wave function is zero. Imposing the nodal constraint can be achieved either by deleting the walkers which attempt to cross the nodes or by using the short-time importance sampled Green's function, where walkers do not cross the nodes in the limit of a zero time step. This procedure is known as the *fixed-node approximation*. Forbidding node crossing is equivalent to finding the exact solution with the boundary condition of having the same nodes as the trial wave function. The Schrödinger equation is therefore solved exactly inside the nodal regions but not at the nodes where the solution will have a discontinuity of the derivatives. The fixed-node solution will be exact only if the nodes of the trial wave function are exact. In general, the fixed-node energy will be an upper bound to the exact energy. A cut through the nodal surface of the N and P atoms for a simple Hartree–Fock and a highly-accurate wave function (Figure 8.5) reveals that considerable differences are possible which are atom dependent and directly translate in a larger size of the fixed-node error for the N atom when a mono-determinantal wave function is used [59].

The fixed-node DMC algorithm can also be used to study excited states. There is no particular difficulty in applying DMC to the lowest state of a given symmetry by simply employing a trial wave function of the proper spatial and spin symmetry.[4] For excited states which are energetically not the

4 More precisely, the DMC energy is variational if the trial function transforms according to a 1-dimensional irreducible representation of the symmetry group of the Hamiltonian [60].

lowest in their symmetry, all that we know is that fixed-node DMC will give the exact solution if we employ a trial wave function with the exact nodes [60]. However, there is no variational principle and one may expect a stronger dependence of the result on the choice of the wave function, which is now not only used to overcome the fermion-sign problem but also to select the state of interest. In our experience, unless we intentionally generate a wave function with a large overlap with the ground-state one, we do not suffer from lack of variationality in the excited-state calculation. In fact, the use of simplistic wave functions (e.g., HOMO-LUMO Hartree–Fock, configuration-interaction singles, non-reoptimized truncated complete-active-space expansions) has been shown to generally lead to an overestimation of the excitation energy also in DMC, especially when the excited state has a strong multi-determinant character [61]. Consequently, while DMC cannot cure the shortcomings of a poor wave function, such a choice will likely yield smaller fixed-node errors in the ground than the excited state and, ultimately, an overestimation of the DMC excitation energy.

8.4 Wave Functions and their Optimization

The key quantity which determines the quality of a VMC and a fixed-node DMC calculation is the trial wave function. The choice of the functional form of the wave function and its optimization within VMC are key steps in a QMC calculation as they are crucial elements to obtain accurate results already at the VMC level and to reduce the fermionic-sign error in a subsequent DMC calculation.

Most QMC studies of electronic systems have employed trial wave functions of the so-called Jastrow–Slater form, that is, the product of a sum of determinants of single-particle orbitals and a Jastrow correlation factor as

$$\Psi = \mathcal{J} \sum_k c_k D_k \,, \tag{8.18}$$

where D_k are Slater determinants of single-particle orbitals and the Jastrow correlation function is a positive function of the interparticle distances, which explicitly depends on the electron–electron separations. The Jastrow factor plays an important role as it is used to impose the Kato cusp conditions and to cancel the divergences in the potential at the inter-particle coalescence points. This leads to a smoother behavior of the local energy and therefore more accurate and efficient VMC as well as DMC calculations thanks to the smaller time-step errors and reduced fluctuations in the branching factor.

Moreover, the Jastrow factor introduces important correlations beyond the short electron–electron distances [62] and QMC wave functions enjoy therefore a more compact determinantal expansion than conventional quantum chemical methods. Even though the positive Jastrow function does not directly alter the nodal structure of the wave function which is solely determined by the antisymmetric part, the optimal determinantal component in a QMC wave function will be different than the one obtained for instance in a multi-configuration self-consistent-field calculation (MCSCF) in the absence of the Jastrow factor. Upon optimization of the QMC wave function, the determinantal component will change and it is often possible to obtain converged energy differences in VMC and DMC with relatively short determinantal expansions in a chosen active space. Furthermore, thanks to the presence of the Jastrow factor, QMC results are generally less dependent on the basis set. For instance, excitations and excited-state gradients show a faster convergence with basis set than multi-configurational approaches, and an augmented double basis set with polarization functions is often sufficient in both VMC and DMC for the description of excited-state properties [24, 63, 64].

An important requirement for the optimization of the many parameters in a QMC wave function is the ability to efficiently evaluate the derivatives of the wave function and the action of the Hamiltonian on these derivatives during a QMC run. In general, this is central to the computation of low-variance estimators of derivatives of the total energy as, for instance, the derivatives with respect to the nuclear coordinates (i.e., interatomic forces). Computing these derivatives at low cost is, therefore, crucial to enable higher accuracy as well as to extend the application of QMC to larger systems and a broader range of molecular properties. Automatic differentiation was successfully applied for the computation of analytical derivatives [12] but the application to large computer codes is not straightforward and the memory requirements might become prohibitive.

Recently, an efficient and simple analytical formulation has been developed to compute a complete set of derivatives of the wave function and of the local energy with the same scaling per Monte Carlo step as computing the energy alone both for single- and multi-determinant wave functions [13, 14]. This formulation relies on the straightforward manipulation of matrices evaluated on the occupied and virtual orbitals and can be very simply illustrated in the case of a single determinant in the absence of a Jastrow factor:

$$D = \det(A) = |\phi_1 \phi_2 \dots \phi_N| \,, \tag{8.19}$$

where A is a Slater matrix defined in terms of the N occupied orbitals, ϕ_i, as $A_{ij} = \phi_j(\mathbf{r}_i)$. For this wave function, it is not difficult to show that the action of a one-body operator $\hat{O} = O(\mathbf{r}_1) + \dots + O(\mathbf{r}_N)$ on the determinant can be written as the trace between the inverse A matrix and an appropriate matrix B,

$$\frac{\hat{O}D}{D} = \mathrm{tr}(A^{-1}B) \,, \tag{8.20}$$

where B is obtained by applying the operator $O(\mathbf{r})$ to the elements of A as

$$B_{ij} = (O\phi_j)(\mathbf{r}_i) \,. \tag{8.21}$$

For instance, if we consider the kinetic operator, we obtain

$$\hat{T}\det(A) = -\frac{1}{2}\sum_i \Delta_i \det(A) = -\frac{1}{2}\sum_i \left[\sum_j \Delta\phi_j(\mathbf{r}_i)\,(A^{-1})_{ji}\det(A)\right] \,, \tag{8.22}$$

which can be rewritten as

$$\frac{\hat{T}\det(A)}{\det(A)} = \sum_i \sum_j B^{\mathrm{kin}}_{ij}\,(A^{-1})_{ji} = \mathrm{tr}(A^{-1}B^{\mathrm{kin}}) \,, \tag{8.23}$$

where $B^{\mathrm{kin}}_{ij} = -\frac{1}{2}\Delta_i A_{ij} = -\frac{1}{2}\Delta\phi_j(\mathbf{r}_i)$. It is possible to show that an equivalent trace expression holds also in the presence of the Jastrow factor but with a B matrix which depends not only on the orbitals but also on the Jastrow factor.

The compact trace expression of a local quantity (Eq. 8.20) offers the advantage that its derivative with respect to a parameter μ can be straightforwardly written as

$$\partial_\mu \frac{\hat{O}D}{D} = \mathrm{tr}(A^{-1}\partial_\mu B - X\partial_\mu A) \,, \tag{8.24}$$

where $\partial_\mu A$ and $\partial_\mu B$ are the matrices of the derivatives of the elements of A and B, respectively, and the matrix X is defined as

$$X = A^{-1}BA^{-1} \,. \tag{8.25}$$

This can easily be derived by using $\partial_\mu(A^{-1}) = -A^{-1}\partial_\mu A\, A^{-1}$ and the cyclic property of the trace. Therefore, if one computes and stores the matrix X, it is then possible to evaluate derivatives at

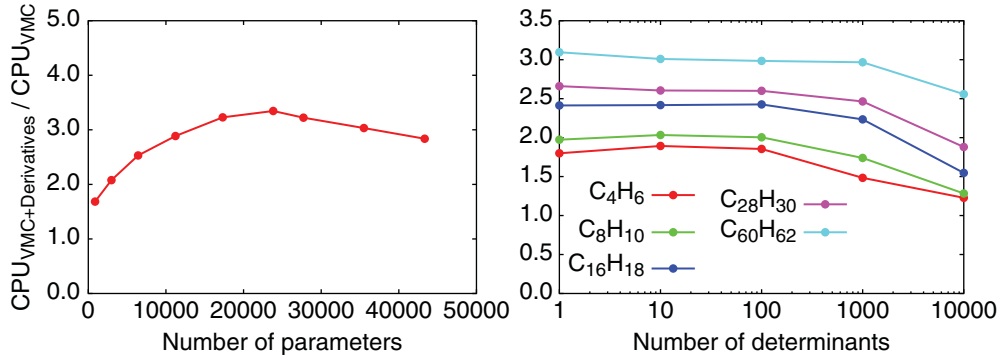


Figure 8.6 Cost per Monte Carlo step of computing a complete set of derivatives of the wave function and local energy relative to a VMC run where only the energy is calculated. Left: increasing number of variational parameters for the series C_nH_{n+2} with n between 4 and 44 (reproduced from Ref. [13]). Right: increasing number of determinants in the Jastrow–Slater wave function for C_nH_{n+2} with n between 4 and 60 (reproduced from Ref. [14]).

the cost of calculating a trace, namely, $O(N^2)$. Consequently, this procedure enables, for instance, the efficient calculation of the $O(N)$ derivatives of the local energy ($\hat{O} = \hat{H}$) with respect to the nuclear coordinates, reducing the scaling of computing the interatomic forces per Monte Carlo step to the one of the energy, namely, $O(N^3)$. The same scaling is also obtained for the optimization of the orbital parameters as shown in Figure 8.6 (left panel) and further discussed in Ref. [13]. This simple formulation and its advantages in the calculation of energy derivatives can be extended to multi-determinant wave functions to achieve a cost in the computation of a set of derivatives proportional to the one of evaluating the energy alone [14] as illustrated in Figure 8.6 (right panel) for the interatomic derivatives.

With all the derivatives of the wave function and the corresponding local quantities at hand, the next step is to use them for the optimization of the wave function. The use of wave functions with a large number of parameters requires efficient algorithms and the two most commonly used approaches, the linear method and the stochastic reconfiguration method, are discussed in the following. We will begin with the simpler case of the optimization in the ground state (or an excited state which is energetically the lowest for a given symmetry).

8.4.1 Stochastic Reconfiguration Method

In the stochastic reconfiguration (SR) method [65, 66], one starts from a given wave function, Ψ, and obtains an improved state by applying the operator $(1 - \tau\hat{H})$, namely, a first-order expansion of the operator $e^{-\tau\hat{H}}$ used in DMC. The new state is then projected in the space spanned by the current wave function and its derivatives, $\{\Psi_i\} = \{\Psi, \partial_i\Psi\}$ as

$$\sum_{j=0}^{N_p} \delta p_j |\Psi_j\rangle = \hat{P}^{\mathrm{SR}}(1 - \tau\hat{H})|\Psi\rangle\,, \tag{8.26}$$

where N_p is the number of wave function parameters and $\Psi_0 = \Psi$. By taking the internal product with $\langle\Psi_i|$ and eliminating the scaling δp_0 through the $i = 0$ equation, one derives a set of equations for $i = 1, \ldots, N_p$, which can be written in matrix notation as

$$\bar{\mathbf{S}}\,\delta\mathbf{p} = -\frac{\tau}{2}\mathbf{g}\,, \tag{8.27}$$

where **g** is the gradient of the energy with respect to the parameters,

$$g_i = 2\left[\frac{\langle\Psi_i|\hat{H}|\Psi\rangle}{\langle\Psi|\Psi\rangle} - E_\mathrm{v}\frac{\langle\Psi_i|\Psi\rangle}{\langle\Psi|\Psi\rangle}\right] \equiv \partial_i E_\mathrm{v}\,, \tag{8.28}$$

and $\bar{\mathbf{S}}$ is related to the overlap matrix between the derivatives, **S**, as

$$\bar{S}_{ij} = \frac{\langle\Psi_i|\Psi_j\rangle}{\langle\Psi|\Psi\rangle} - \frac{\langle\Psi_i|\Psi\rangle}{\langle\Psi|\Psi\rangle}\frac{\langle\Psi|\Psi_j\rangle}{\langle\Psi|\Psi\rangle} \equiv S_{ij} - S_{i0}S_{0j}\,. \tag{8.29}$$

With an appropriate choice of τ, a new set of parameters can be determined as $p_i' = p_i + \delta p_i$ and the procedure iterated until convergence. Therefore, the SR method is like a Newton approach where one follows the downhill gradient of the energy, but using the matrix $\bar{\mathbf{S}}$ instead of the Hessian of the energy. Even though the method can display a slow convergence since τ scales like the inverse of the energy range spanned by the wave function derivatives [67], it was recently employed to successfully optimize very large numbers of parameters [14, 15]. We will come back to this point when discussing the linear method below.

In a VMC run of SR optimization, one needs to compute the gradient and the overlap matrix **S** by sampling the distribution ρ given by the square of the current wave function (Eq. 8.2) as

$$\begin{aligned} S_{ij} &= \frac{\langle\Psi_i|\Psi_j\rangle}{\langle\Psi|\Psi\rangle} = \frac{\int \Psi_i(\mathbf{R})\Psi_j(\mathbf{R})\mathrm{d}\mathbf{R}}{\int \Psi(\mathbf{R})^2\mathrm{d}\mathbf{R}} \\ &= \int \frac{\Psi_i(\mathbf{R})}{\Psi(\mathbf{R})}\frac{\Psi_j(\mathbf{R})}{\Psi(\mathbf{R})}\rho(\mathbf{R})\mathrm{d}\mathbf{R} \approx \frac{1}{M}\sum_k^M \frac{\Psi_i(\mathbf{R}_k)}{\Psi(\mathbf{R}_k)}\frac{\Psi_j(\mathbf{R}_k)}{\Psi(\mathbf{R}_k)}\,. \end{aligned} \tag{8.30}$$

For a large number of parameters, not only does the storage of this matrix become problematic but also its calculation whose cost scales as $\mathcal{O}(MN_p^2)$. However, if we use a conjugate gradient method to solve the linear equations (8.27), we only need to repeatedly evaluate **S** acting on a trial vector of parameter variations as

$$\sum_{j=1}^{N_p} S_{ij}\delta p_j = \frac{1}{M}\sum_k^M \frac{\Psi_i(\mathbf{R}_k)}{\Psi(\mathbf{R}_k)}\sum_{j=1}^{N_p}\frac{\Psi_j(\mathbf{R}_k)}{\Psi(\mathbf{R}_k)}\,\delta p_j\,, \tag{8.31}$$

where the order of the sums in the last expression has been swapped [68]. Therefore, if we compute and store the $M \times N_p$ matrix of the ratios $\Psi_i(\mathbf{R}_k)/\Psi(\mathbf{R}_k)$ during the Monte Carlo run, we can reduce the memory requirements by exploiting the intrinsic parallelism of Monte Carlo simulations: we can employ a small M per core and increase instead the number of cores to obtain the desired statistical accuracy. The computational cost of solving the SR equations is also reduced to $\mathcal{O}(N_\mathrm{CG}MN_p)$, where N_CG is the number of conjugate gradient steps, which we have found to be several orders of magnitude smaller than the number of parameters in recent optimization of large determinantal expansions [14, 15].

8.4.2 Linear Method

The linear optimization method is related to the so-called super configuration interaction (super-CI) approach used in quantum chemistry to optimize the orbital parameters in a multi-determinant wave function. The starting point is the normalized wave function,

$$|\bar{\Psi}\rangle = \frac{1}{\sqrt{\langle\Psi|\Psi\rangle}}|\Psi\rangle\,, \tag{8.32}$$

which we expand to first order in the parameter variations around the current values in the basis of the current wave function and its derivatives, $\{\bar{\Psi}_i\} = \{\bar{\Psi}, \partial_i\bar{\Psi}\}$ with $\bar{\Psi}_0 = \bar{\Psi}$. The important

advantage of working with the "barred" functions is that they are orthogonal to the current wave function since

$$|\bar{\Psi}_i\rangle = \frac{1}{\sqrt{\langle\Psi|\Psi\rangle}}\left(|\Psi_i\rangle - \frac{\langle\Psi|\Psi_i\rangle}{\langle\Psi|\Psi\rangle}|\Psi\rangle\right), \tag{8.33}$$

which has been found to yield better (non-linear) parameter variations and a more robust optimization than simply using the derivatives of the wave function.

The change of the parameters $\delta\mathbf{p}$ is then determined by minimizing the expectation value of the Hamiltonian on the linearized wave function in the basis $\{\bar{\Psi}_i\}$, which leads to the generalized eigenvalue equations:

$$\sum_{j=0}^{N_p} \bar{H}_{ij}\delta p_j = E_{\text{lin}} \sum_{j=0}^{N_p} \bar{S}_{ij}\delta p_j\,, \tag{8.34}$$

where $\bar{H}_{ij} = \langle\bar{\Psi}_i|\hat{H}|\bar{\Psi}_j\rangle$ and $\bar{S}_{ij} = \langle\bar{\Psi}_i|\bar{\Psi}_j\rangle$. We note that the overlap $\bar{\mathbf{S}}$ is equivalent to the expression introduced above in the SR scheme (Eq. 8.29). A new set of parameters can be generated as $p_i' = p_i + \delta p_i/\delta p_0$ and the algorithm iterated until convergence. Importantly, in a Monte Carlo run, the matrix $\bar{\mathbf{H}}$ will not be symmetric for a finite sample and a non-obvious finding is that the method greatly benefits from reduced fluctuations if one does *not* symmetrize the Hamiltonian matrix, as originally shown by Nightingale and Melik-Alaverdian for the optimization of the linear parameters [32]. Other important modifications can be introduced to further stabilize the approach and improve the convergence as discussed in Ref. [7].

It is simple to recognize that, at convergence, the linear method leads to an optimal energy if we express explicitly the secular equations above in matrix form as

$$\begin{pmatrix} E_{\text{v}} & \frac{1}{2}g^{\text{T}} \\ \frac{1}{2}g & \bar{\mathbf{H}} \end{pmatrix}\begin{pmatrix} \delta p_0 \\ \delta\mathbf{p} \end{pmatrix} = E_{\text{lin}}\begin{pmatrix} 1 & 0 \\ 0 & \bar{\mathbf{S}} \end{pmatrix}\begin{pmatrix} \delta p_0 \\ \delta\mathbf{p} \end{pmatrix}, \tag{8.35}$$

where we have used that $\bar{H}_{00} = \langle\bar{\Psi}|\hat{H}|\bar{\Psi}\rangle$ is the current energy and the elements of the first column and row, $\bar{H}_{i0} = \langle\bar{\Psi}_i|\hat{H}|\bar{\Psi}\rangle$ and $\bar{H}_{0i} = \langle\bar{\Psi}|\hat{H}|\bar{\Psi}_i\rangle$, respectively, are both mathematically equal to the components of the energy gradient (Eq. 8.28). Therefore, when the wave function parameters are optimal, the variations with respect to the current wave function will no longer couple to it ($\delta p_i = 0$) and the $\bar{H}_{i0}$ and $\bar{H}_{0i}$ elements must therefore become zero. This directly implies that the gradient of the energy with respect to the parameters is identically zero.

To further understand how the linear method is related to other optimization schemes, one can recast its equations as a Newton method [69]:

$$(\mathbf{A} + \alpha\bar{\mathbf{S}})\delta\mathbf{p} = -\frac{1}{2}\mathbf{g}\,, \tag{8.36}$$

where $\mathbf{A} = \bar{\mathbf{H}} - E_{\text{v}}\bar{\mathbf{S}}$ and $\alpha = E_{\text{v}} - E_{\text{lin}} > 0$. Therefore, the parameters are varied along the downhill gradient of the energy with the use of an approximate Hessian $\mathbf{A}$ level-shifted by the positive definite matrix $\alpha\bar{\mathbf{S}}$. The presence of the latter renders the optimization more stable and effective than the actual Newton method even when the exact Hessian matrix is used. While the linear method is in principle significantly more efficient than, for instance, the SR approach, we find that its stochastic realization suffers from large fluctuations in the elements of $\bar{\mathbf{H}}$ when one optimizes the orbital parameters or the linear coefficients of particularly extended multi-determinant wave functions (where the derivatives are very different from the actual wave function used in the sampling). As a result, the optimization requires long VMC runs to achieve reliable variations in the parameters or a large shift added to the diagonal elements of $\bar{\mathbf{H}}$ (except $\bar{H}_{00}$) [7] to stabilize the procedure. In

these cases, we find that the SR scheme, which only makes use of the $\bar{\mathbf{S}}$ matrix, is more robust and efficient since it allows less strict requirements on the error bars.

Finally, we note that, as in the SR scheme, it is possible to avoid explicitly constructing the full matrices $\bar{\mathbf{H}}$ and $\bar{\mathbf{S}}$: one stores the local quantities $\Psi_i(\mathbf{R})/\Psi(\mathbf{R})$ and $\hat{H}\Psi_i(\mathbf{R})/\Psi(\mathbf{R})$ in the Monte Carlo run and uses for instance a generalized Davidson algorithm to find the eigenvectors where only matrix-vector products with trial vectors are evaluated, significantly reducing the computational and memory requirements [68].

8.5 Excited States

For excited states of a different symmetry than the ground state, one can construct a trial wave function of the desired space and spin symmetry (with an appropriate choice of the determinantal component) and apply either the SR or the linear method to minimize the energy in VMC, subsequently refining the calculation in DMC. For excited states which are energetically not the lowest in their symmetry class, one can instead follow different routes as in other quantum chemistry methods to find an accurate excited-state wave function. We will begin by describing the possibilities within energy minimization and then consider optimization schemes targeting the variance of the energy which has a minimum for each eigenstate of the Hamiltonian.

8.5.1 Energy-Based Methods

While the linear method is generally employed for ground-state wave function optimizations, it is in fact possible to use it in a state-specific manner for the optimization of excited states [70]. One can target a higher-energy state and linearize the problem with respect to the chosen state. Solving a generalized eigenvalue problem as in Eq. (8.35) will yield lower energy roots as well as the state of interest. The resulting wave function will be only approximately orthogonal to the lower ones since orthogonality is only imposed in the basis of the variations of the optimal target wave function with respect to the parameters. Furthermore, since following such a higher root leads to the optimization of a saddle point in parameter space, the procedure may exhibit convergence problems so that the parameters do not converge to the desired state. One may also observe flipping of the roots: As the optimization proceeds, the optimized excited target state can obtain a lower eigenvalue than the unoptimized ground state. Such problems will be particularly severe in the case of close degeneracy as in proximity of conical intersection regions.

A different route to optimize multiple states of the same symmetry lies in the generalization of state-average (SA) approaches to QMC [9]. We start from a set of Jastrow–Slater wave functions for the multiple states that are constructed as linear combinations of determinants multiplied by a Jastrow factor as

$$\Psi^I = \mathcal{J} \sum_k c_k^I D_k \,, \tag{8.37}$$

where the index I labels the states. The wave functions of the different states are therefore characterized by different linear coefficients c_i^I but share a common set of orbitals and the Jastrow factor $\mathcal{J}$.

The optimal linear coefficients c_i^I can be easily determined through the solution of the generalized eigenvalue equations

$$\sum_j H_{ij} c_j^I = E_I \sum_j S_{ij} c_j^I \,, \tag{8.38}$$

where the matrix elements are here given by

$$H_{ij} = \langle \mathcal{J}D_i | \hat{H} | \mathcal{J}D_j \rangle \quad \text{and} \quad S_{ij} = \langle \mathcal{J}D_i | \mathcal{J}D_j \rangle \,, \tag{8.39}$$

and are computed in a VMC run, where we do not symmetrize the Hamiltonian matrix for finite Monte Carlo sampling to reduce the fluctuations of the parameters as discussed above for the general linear method. After diagonalization of Eq. (8.38), the optimal linear coefficients are obtained and, at the same time, orthogonality between the individual states is automatically enforced.

To obtain a robust estimate of the linear coefficients of multiple states, it is important that the distribution sampled to evaluate H_{ij} and S_{ij} has a large overlap with all states of interest (and all lower lying states). A suitable guiding wave function can for instance be constructed as

$$\Psi_g = \sqrt{\sum_I |\Psi_I|^2} \,, \tag{8.40}$$

and the distribution ρ_g in the VMC run defined as the square of this guiding function. The matrix elements S_{ij} (and, similarly, H_{ij}) are then evaluated in the Monte Carlo run as

$$\begin{aligned} \frac{S_{ij}}{\langle \Psi_g | \Psi_g \rangle} &= \frac{\int \mathcal{J}D_i(\mathbf{R}) \mathcal{J}D_j(\mathbf{R}) \mathrm{d}\mathbf{R}}{\int \Psi_g^2(\mathbf{R}) \mathrm{d}\mathbf{R}} = \int \frac{\mathcal{J}D_i(\mathbf{R})}{\Psi_g(\mathbf{R})} \frac{\mathcal{J}D_j(\mathbf{R})}{\Psi_g(\mathbf{R})} \rho_g(\mathbf{R}) \mathrm{d}\mathbf{R} \\ &\approx \frac{1}{M} \sum_k^M \frac{\mathcal{J}D_i(\mathbf{R}_k)}{\Psi_g(\mathbf{R}_k)} \frac{\mathcal{J}D_j(\mathbf{R}_k)}{\Psi_g(\mathbf{R}_k)} \,. \end{aligned} \tag{8.41}$$

We note that we can introduce the denominator $\langle \Psi_g | \Psi_g \rangle$ if we simply divide by it both sides of Eq. (8.38).

As done in state-average multi-configurational approaches to obtain a balanced description of the states of interest, one can optimize the non-linear parameters of the orbitals and the Jastrow factor by minimizing the state-average energy

$$E^{\mathrm{SA}} = \sum_I w_I \frac{\langle \Psi^I | \hat{H} | \Psi^I \rangle}{\langle \Psi^I | \Psi^I \rangle} \,, \tag{8.42}$$

with the weights of the states w_I kept fixed and $\sum_I w_I = 1$. The gradient of the SA energy can be rewritten as

$$g_i^{\mathrm{SA}} = \sum_I w_I \langle \bar{\Psi}_i^I | \hat{H} | \bar{\Psi}^I \rangle \,, \tag{8.43}$$

where, similar to Eq. (8.33), we have introduced for each state the variations, $|\Psi_i^I\rangle = |\partial_i \Psi^I\rangle$, and the corresponding "barred" functions orthogonal to the current state $|\Psi^I\rangle$:

$$|\bar{\Psi}_i^I\rangle = \frac{1}{\sqrt{\langle \Psi^I | \Psi^I \rangle}} \left(|\Psi_i^I\rangle - \frac{\langle \Psi^I | \Psi_i^I \rangle}{\langle \Psi^I | \Psi^I \rangle} |\Psi^I\rangle \right) . \tag{8.44}$$

The variations in the parameters can be obtained as the lowest-energy solution of the generalized eigenvalue equation in analogy to the linear method for the ground state,

$$\begin{pmatrix} E^{\mathrm{SA}} & \frac{1}{2}(g^{\mathrm{SA}})^{\mathrm{T}} \\ \frac{1}{2} g^{\mathrm{SA}} & \bar{\mathbf{H}}^{\mathrm{SA}} \end{pmatrix} \begin{pmatrix} \delta p_0 \\ \delta \mathbf{p} \end{pmatrix} = E \begin{pmatrix} 1 & 0 \\ 0 & \bar{\mathbf{S}}^{\mathrm{SA}} \end{pmatrix} \begin{pmatrix} \delta p_0 \\ \delta \mathbf{p} \end{pmatrix} . \tag{8.45}$$

The state-average matrix elements are defined as

$$\bar{H}_{ij}^{\mathrm{SA}} = \sum_I w_I \frac{\langle \bar{\Psi}_i^I | \hat{H} | \bar{\Psi}_j^I \rangle}{\langle \bar{\Psi}^I | \bar{\Psi}^I \rangle} \,, \tag{8.46}$$

and an analogous expression for $\bar{S}_{ij}^{\mathrm{SA}}$ is introduced. To compute these matrix elements in VMC, we perform a single run sampling the square of a guiding wave function Ψ_g (Eq. 8.40) and compute the numerators and denominators in the matrix expressions for all relevant states. We note that the state-average equations (Eq. 8.45) are not obtained by minimizing a linearized expression of the SA energy (Eq. 8.42) but are simply inspired by the single-state case. However, since the first row and column in Eq. (8.45) are given by the gradient of the SA energy, at convergence, the optimal parameters minimize the SA energy. We find that the use of these state-average Hamiltonian and overlap matrices leads to a similar convergence behavior as the linear method for a single state.

Following this procedure, the algorithm alternates between the minimization of the linear and the non-linear parameters until convergence is reached. The obtained energy E_{SA} is stationary with respect to variations of all parameters while the energies of the individual states, E_I, are only stationary with respect to the linear but not the orbital and Jastrow parameters. If the ground state and the target excited state should be described by very different orbitals, a state-specific approach may yield more accurate energies.

8.5.2 Time-Dependent Linear-Response VMC

A very different approach to the computation of multiple excited states is a VMC formulation of linear-response theory [71]. Given a starting wave function Ψ with optimal parameters $\mathbf{p}_0$, a time-dependent perturbation $\hat{V}(t)$ is introduced in the Hamiltonian $\hat{H}$ with the coupling constant γ as

$$\hat{H}(t) = \hat{H} + \gamma \hat{V}(t)\,, \tag{8.47}$$

so that the ground-state wave function itself becomes time-dependent as the variational parameters $\mathbf{p}(t)$ evolve in time. It is convenient to work with a wave function subject to an intermediate normalization,

$$|\bar{\Psi}(t)\rangle = \frac{|\Psi(t)\rangle}{\langle \Psi_0|\Psi(t)\rangle}\,, \tag{8.48}$$

where the starting wave function $\Psi_0 \equiv \Psi$ is taken to be normalized. This choice leads to wave function variations to first and second order that are orthogonal to the current optimal wave function Ψ_0.

At each time t, one can apply the Dirac–Frenkel variational principle to obtain the parameters $\mathbf{p}(t)$ as

$$\frac{\partial}{\partial p_i^*} \frac{\langle \bar{\Psi}(t)|\hat{H}(t) - i\frac{\partial}{\partial t}|\bar{\Psi}(t)\rangle}{\langle \bar{\Psi}(t)|\bar{\Psi}(t)\rangle} = 0, \tag{8.49}$$

where the parameters can now in general be complex. To apply this principle to linear order in γ, the wave function is expanded to second order in $\delta\mathbf{p}(t)$ around $\mathbf{p}_0$:

$$|\bar{\Psi}(t)\rangle = |\bar{\Psi}_0\rangle + \sum_i \delta p_i(t)|\bar{\Psi}_i\rangle + \frac{1}{2}\sum_{ij} \delta p_i(t)\delta p_j(t)|\bar{\Psi}_{ij}\rangle\,, \tag{8.50}$$

with $|\bar{\Psi}_i\rangle = |\partial_i \bar{\Psi}\rangle$ and $|\bar{\Psi}_{ij}\rangle = |\partial_i \partial_j \bar{\Psi}\rangle$ computed at the parameters $\mathbf{p}_0$. These can be explicitly written as

$$\begin{aligned}
|\bar{\Psi}_i\rangle &= |\Psi_i\rangle - \langle \Psi_0|\Psi_i\rangle|\Psi_0\rangle \\
|\bar{\Psi}_{ij}\rangle &= |\Psi_{ij}\rangle - \langle \Psi_0|\Psi_i\rangle|\Psi_j\rangle - \langle \Psi_0|\Psi_j\rangle|\Psi_i\rangle \\
&\quad + (2\langle \Psi_0|\Psi_i\rangle\langle \Psi_0|\Psi_j\rangle - \langle \Psi_0|\Psi_{ij}\rangle)|\Psi_0\rangle\,,
\end{aligned} \tag{8.51}$$

where we use the same notation as above for the derivatives of Ψ, namely, $|\Psi_i\rangle = |\partial_i \Psi\rangle$ and $|\Psi_{ij}\rangle = |\partial_i \partial_j \Psi\rangle$. Inserting this wave function in Eq. (8.49) and keeping only the first order terms in $\delta\mathbf{p}(t)$, in the limit of $\gamma \to 0$, one obtains

$$\mathbf{A}\ \delta\mathbf{p}(t) + \mathbf{B}\ \delta\mathbf{p}(t)^* = i\bar{\mathbf{S}}\ \frac{\partial \delta\mathbf{p}(t)}{\partial t}\,, \tag{8.52}$$

with the matrix elements $A_{ij} = \langle \bar{\Psi}_i | \hat{H} - E_0 | \bar{\Psi}_j \rangle = \bar{H}_{ij} - E_0 \bar{S}_{ij}$ and $B_{ij} = \langle \bar{\Psi}_{ij} | \hat{H} | \Psi_0 \rangle$. If we search for an oscillatory solution,

$$\delta\mathbf{p}(t) = \mathbf{X}_n e^{-i\omega_n t} + \mathbf{Y}_n^* e^{i\omega_n t}\,, \tag{8.53}$$

with ω_n an excitation energy and $\mathbf{X}_n$ and $\mathbf{Y}_n$ the response vectors, we obtain the well-known linear-response equations, here formulated as a non-Hermitian generalized eigenvalue equation,

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{pmatrix} \begin{pmatrix} \mathbf{X}_n \\ \mathbf{Y}_n \end{pmatrix} = \omega_n \begin{pmatrix} \bar{\mathbf{S}} & \mathbf{0} \\ \mathbf{0} & -\bar{\mathbf{S}}^* \end{pmatrix} \begin{pmatrix} \mathbf{X}_n \\ \mathbf{Y}_n \end{pmatrix}\,. \tag{8.54}$$

Neglecting $\mathbf{B}$ leads to the Tamm–Dancoff approximation,

$$\mathbf{A}\mathbf{X}_n = \omega_n \bar{\mathbf{S}} \mathbf{X}_n\,, \tag{8.55}$$

which is equivalent to the generalized eigenvalue equations of the linear method (Eq. 8.35) for an optimized ground-state wave function (when the gradients of the energy are therefore zero):

$$\bar{\mathbf{H}}\mathbf{X}_n = (\omega_n + E_0)\bar{\mathbf{S}}\mathbf{X}_n\,. \tag{8.56}$$

The energy E_0 is the variational energy, E_v, for the optimized ground state. Therefore, upon optimization of the ground-state wave function in the linear method, we can simply use the higher roots resulting from the diagonalization of the equations to estimate the excitation energies as $\omega_n = (E_\mathrm{lin})_n - E_0$ together with the oscillator strengths [71].

The time-dependent linear-response VMC approach has so far only been applied to the excitations of the beryllium atom within the Tamm–Dancoff approximation and with a simple single-determinant Jastrow–Slater wave function [71]. These calculations represent an interesting proof of principle that multiple excitations of different space and spin symmetry can be readily obtained after optimizing the ground-state wave function. A systematic investigation with multi-configurational wave functions is needed to fully access the quality of the approach, also beyond the Tamm–Dancoff approximation.

8.5.3 Variance-Based Methods

Variance minimization is a different approach to optimize the wave function compared to the methods described so far, which, in principle, allows optimization of excited states in a state-specific fashion. The target quantity for the minimization is the variance of the energy:

$$\sigma_\mathrm{v}^2 = \frac{\langle \Psi | (\hat{H} - E_\mathrm{v})^2 | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{\int \Psi^2(\mathbf{R})(E_\mathrm{L}(\mathbf{R}) - E_\mathrm{v})^2 \mathrm{d}\mathbf{R}}{\int \Psi^2(\mathbf{R}) \mathrm{d}\mathbf{R}}\,. \tag{8.57}$$

While the (global) minimum of the variational energy is only obtained for the ground state, the variance has a known minimum of zero for each eigenstate of the Hamiltonian. The optimization of the variance can be performed using either a Newton approach with an approximate expression of the Hessian of the variance [5] or reformulated as a generalized eigenvalue problem, namely, a linear method for the optimization of the variance [69]. For excited states, the initial guess of the

trial wave function might, however, be very important to select a specific state and ensure that the minimization of the variance leads to the correct local minimum.

More robust state-specific variational principles for excited states can be formulated so that the optimization of the excited state yields a minimum close to an initial target energy. A simple possibility is to substitute the wave function-dependent average energy in σ_v with a guess value ω as in

$$\sigma_\omega^2 = \frac{\langle\Psi|(\hat{H}-\omega)^2|\Psi\rangle}{\langle\Psi|\Psi\rangle} \equiv (E_v-\omega)^2 + \sigma_v^2 \,, \tag{8.58}$$

as was done in the early applications of variance minimization on a fixed Monte Carlo sample [72], where ω was chosen equal to a target value at the beginning of the optimization and then adjusted to the current best energy. If one updates ω in this manner, minimizing σ_ω is equivalent to minimizing σ_v.

Alternatively, minimization of the variance can also be achieved by optimizing the recently proposed functional Ω defined as

$$\Omega = \frac{\langle\Psi|\omega-\hat{H}|\Psi\rangle}{\langle\Psi|(\omega-\hat{H})^2|\Psi\rangle} = \frac{\omega-E_v}{(\omega-E_v)^2+\sigma_v} \,, \tag{8.59}$$

where ω is adjusted during the optimization to be equal to the current value of $E_v - \sigma_v$ [10, 11]. While the functional has formally a minimum for a state with energy directly above $\omega + \sigma_v$, keeping ω fixed would lead to lack of size consistency in the variational principle [11]. Therefore, after some initial iterations, ω is gradually varied to match the current value of $E_v - \sigma_v$ required to achieve variance minimization. As in the case of the energy and the variance, this functional can be optimized in VMC through a generalization of the linear method [10, 30].

8.6 Applications to Excited States of Molecular Systems

The status of excited-state quantum Monte Carlo calculations closely parallels the methodological developments that have characterized the last decade as we have outlined above in the context of wave function optimization. Since the early applications to excited states, QMC methods were mainly employed as a tool to compute vertical excitation energies and validate results of more approximate methods. Input from other – sometimes much less accurate – quantum chemical approaches was, however, then used for the construction of the wave function, whose determinant component was generally not optimized in the presence of the Jastrow factor. The success of the calculation was therefore often heavily relying on the ability of DMC to overcome possible shortcomings of the chosen trial wave function. This must be contrasted to the recent situation of VMC having matured to a fully self-consistent method as regards the wave function *and* the geometry with a rich ecosystem of tools ranging from basis sets and pseudopotentials to multi-scale formulations.

One of the first QMC computations of two states of the same symmetry was carried out for the H_2 molecule [73]: the wave function was obtained from a multi-reference calculation and DMC was shown to be able, in this case, to correct for the wave function bias. Over the subsequent years, a number of studies of vertical excitation energies were carried out with this basic recipe, namely, performing DMC calculations on a given simple wave function obtained at a lower level of theory [74–83]. Some of these early excited-state calculations were in fact pioneering as they were applied to remarkably large systems such as silicon and carbon nanoclusters with more than 100

atoms [74, 76, 79]. Given the size of the systems, the choice of excited-state wave function was then very simple and consisted of a single determinant correlated with a Jastrow factor and constructed with the HOMO and LUMO orbitals from a density functional theory (DFT) calculation. Nevertheless, the resulting DMC excitation energies clearly represented an improvement on the time-dependent DFT values and captured much of the qualitative physics of the problem. Even though doubts on the validity of this simplistic recipe [82, 83] led researchers to investigate the use of orbitals and pseudopotentials obtained with different density functionals [79] as well as a multi-determinant description [83], a rather heuristic approach erring on the side of computational saving characterized excited-state calculations in this earlier period.

More recently, the development of algorithms for wave function optimizations in a state-specific or state-average fashion has allowed us to better understand the proper ingredients in an excited-state QMC calculation through the study of simple but challenging molecules. In particular, it has become apparent that large improvements in the accuracy of both VMC and DMC excited states can be achieved by optimizing the determinantal component in the presence of the Jastrow factor at the VMC level [9, 61, 64, 67, 70, 84]. For instance, accurate excitations for low-lying states of ethene can only be obtained if the orbitals derived from a complete-active-space self-consistent-field (CASSCF) calculation are reoptimized in the presence of the Jastrow factor to remove spurious valence-Rydberg mixing in the final DMC energies [67]. The analysis of various small organic molecules reveals that the use of simplistic wave functions such as a HOMO-LUMO Hartree–Fock or a configuration-interaction-singles ansatz may lead to significant errors also in DMC [61]. An investigation of the ground and excited states of methylene shows that the optimization of all variational parameters reduces the dependency of both VMC and DMC on the size of the active space employed for the trial wave function as shown in Figure 8.7 [70]. In general, while a minimal requirement is to optimize the linear coefficients together with the Jastrow factor, the optimization of the orbitals is highly recommended, especially if one employs a truncated expansion in computing the excitation energies. Furthermore, evidence has been given that the optimization of excited-state geometries requires the optimization of all wave function parameters in order to obtain accurate gradients and, consequently, geometries [63]. Finally, by construction, linear-response VMC depends strongly on the quality of the ground-state wave function for the

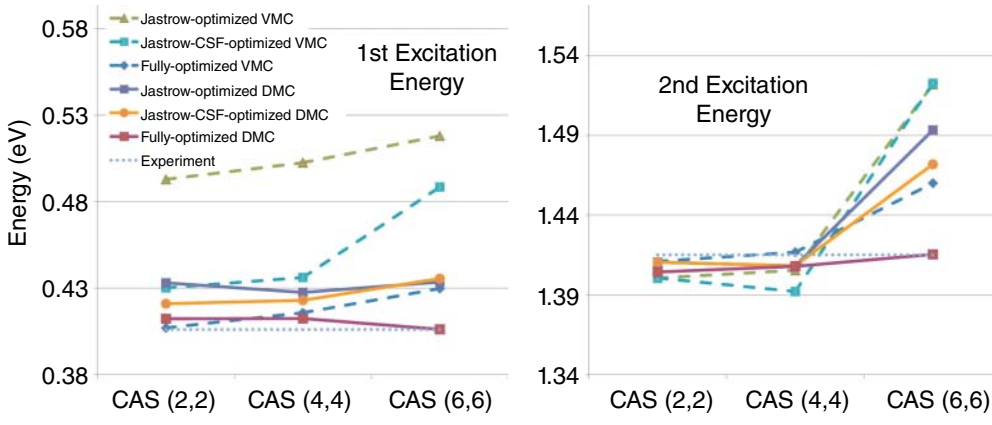


Figure 8.7 Convergence of VMC and DMC adiabatic excitation energies for the first and second excited states of methylene with increasing CAS size. Three levels of optimization have been used for the wave function: Jastrow, Jastrow and linear coefficients, and all parameters. Adapted from Ref. [70].

description of excited states and, therefore benefits considerably from orbital optimization in the ground state [71].

While these and other examples of excited-state QMC calculations clearly illustrate the importance of using wave functions with an adequate description of static correlation and consistently optimized in VMC, they also demonstrate the robustness of QMC approaches and some of their advantages with respect to standard multi-configurational methods. In particular, the VMC and DMC excitations are well converged already when very few determinants of a CAS expansion are kept in the determinantal component of the wave function [9, 84]. Furthermore, the demands on the size of the basis set are also less severe and one can obtain converged excitation energies with rather small basis sets [24, 63, 64]. We note that most of the recent QMC calculations for excited states have attempted to achieve a balanced static description of the states of interest either by employing a CAS in the determinantal component or a truncated multi-reference ansatz where one keeps the union of the configuration state functions resulting from an appropriate truncation scheme (e.g., the sum of the squared coefficients being similar for all states). Interestingly, matching the variance of the states has recently been put forward as a more robust approach to achieve a balanced treatment of the states in the computation of excitation energies [30, 85].

The ability to optimize geometries even in the ground states has been a very recent achievement for QMC methods, so most QMC calculations also outside the Franck–Condon region have been performed on geometries obtained at a different level of theory [30, 61, 70, 83, 86–88]. Nevertheless, these investigations have led to very promising results, showing interesting prospects for the application of QMC to geometry relaxations in the excited state, where most quantum chemical methods either lack the required accuracy or are computationally prohibitive due to their scaling with system size. For example, QMC was successfully employed to assess the accuracy of various time-dependent DFT methods in describing the photochemisty of oxirane through exploration of multiple excited-state potential energy surfaces, also in proximity of conical intersection regions [86, 87]. Another application demonstrating the very good performance of DMC was the study of different conformers of azobenzene in the ground and excited states [88]. To the best of our knowledge, to date, the few attempts to optimize an excited-state geometry via QMC gradients are our studies of the retinal protonated Schiff base model [24, 63] and benchmark calculations on small organic molecules in the gas phase [89] and in a polarizable continuum model [46]. As shown for the retinal minimal model in Figure 8.8, the results are very encouraging as they demonstrate that the QMC structures relaxed in the excited state are in very good agreement with other highly-correlated approaches. As already mentioned, the VMC gradients are sensitive to the quality of the wave function and the orbitals must be reoptimized in VMC to obtain accurate results. Tests also indicate that the use of DMC gradients is not necessary as DMC cannot compensate for the use of an inaccurate wave function while it yields comparable results to VMC when the fully optimized wave function is employed.

Finally, we mention the recent developments of multi-scale methods in combination with QMC calculations for excited states. Multi-scale approaches are particularly relevant for the description of photoactive processes that can be traced back to a region with a limited number of atoms, examples being a chromophore in a protein or a solute in a solvent. While this locality enables us to treat the photoactive region quantum mechanically, excited-state properties can be especially sensitive to the environment (e.g., the polarity of a solvent or nearby residues of a protein), which cannot therefore be neglected but are often treated at lower level of theory. Multi-scale approaches are well-established in traditional quantum chemistry but represent a relatively new area of research in the context of QMC. The first steps in this direction were made by combining VMC with a continuum solvent model, namely, the polarizable continuum model (PCM) [90]. The approach was

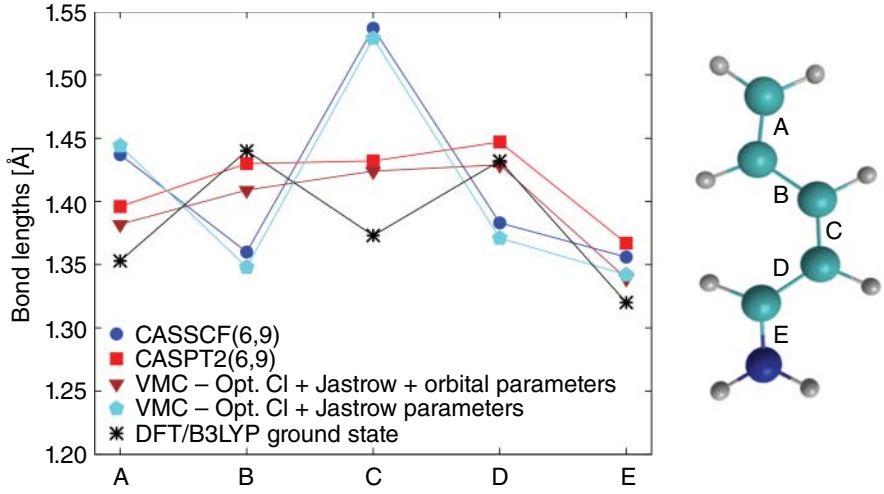


Figure 8.8 Planar excited-state geometry of a retinal protonated Schiff base model optimized with CASSCF, second order perturbation theory (CASPT2), and VMC. Adapted from Ref. [63].

used to investigate solvent effects on the vertical excitation energies of acrolein [44] and on the optimal excited-state geometries of a number of small organic molecules [46]. A notable advantage of VMC/PCM is that the interaction between the polarizable embedding and the solute is described self-consistently at the same level of theory. This stands for instance in contrast to perturbation approaches which include the interaction with the environment obtained self-consistently only at the zero-order level (e.g., CASSCF).

To achieve a more realistic description of the environment, a static molecular mechanics environment coupled via electrostatic interactions with the VMC or DMC chromophore was used to describe the absorption properties of the green fluorescent protein and rhodopsin [91, 92]. The limitations of such a non-polarizable embedding scheme led to further developments, replacing

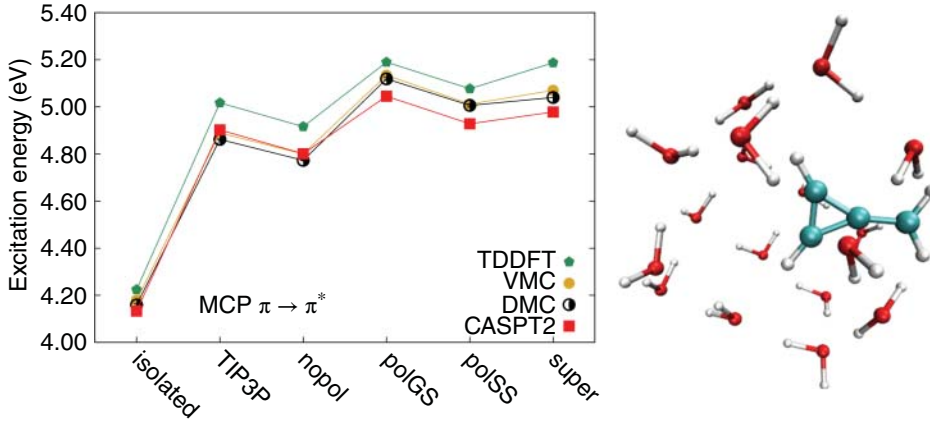


Figure 8.9 Energies of the $\pi \rightarrow \pi^*$ excitation of methylenecyclopropene in vacuum and embedded in a cluster of water molecules, computed with time-dependent DFT, CASPT2, VMC, and DMC. The water molecules are described with a static TIP3P force field and a MMpol approach with no polarization (nopol), with induced ground-state (polGS) and state-specific (polSS) dipoles, and in a supermolecular calculation (super). Adapted from Ref. [47].

the static description with a polarizable one in a so-called VMC/MMpol approach [47]. The force field consists of static partial charges located at the positions of the atoms as well as atomic polarizibilities. These are used to compute induced dipoles in equilibrium with the embedded system at the level of VMC, which are then kept fixed in subsequent DMC calculations. The computation of the QMC excitation energies can be carried out for two polarization schemes: either the induced dipoles are determined for the ground state and used also for the excited state (polGS), or the excitation energy is computed in a state-specific manner as the difference between the ground- and excited-state energies both obtained self-consistently in equilibrium with the respective induced dipoles (polSS). As illustrated in Figure 8.9, the vertical excitation energies of small molecules in water clusters depend strongly on the sophistication of the embedding scheme. Only the polarizable force field with two separate sets of induced dipoles for the ground and excited states leads to a very good agreement with the supermolecular excitation. Again, the QMC results agree with complete-active-space second-order perturbation theory (CASPT2) which however is found to be rather sensitive to the choice of the active space. The most sophisticated QMC embedding scheme has so far been realized using a wave-function-in-DFT method and including differential polarization effects through state-specific embedding potentials [45]. As in the case of the polSS approach, the use of different potentials in the ground and excited state is particularly important for excitations which involve large polarization effects due to a considerable rearrangement of the electron density upon excitation.

8.7 Alternatives to Diffusion Monte Carlo

In some of the applications we presented, VMC has been shown to be sufficient to provide accurate excited-state properties without the need to perform a DMC calculation. The reason is that the burden and complexity of the problem have now been moved from the DMC projection to the construction and optimization at the VMC level of sophisticated wave functions with many parameters. It is therefore natural to ask if there are alternatives to DMC, which do not require us to build and optimize complicated many-body wave functions.

The Cerperley–Bernu method [93] can, in principle, be used to compute the lowest-energy eigenstates and the corresponding relevant matrix elements by constructing a set of many-body basis states and improving upon them through the application of the imaginary-time projection operator also used in DMC. The Hamiltonian and overlap matrices are computed on these improved basis states during the projection and the eigenvalues and eigenstates are then obtained by solving this generalized eigenvalue problem. The method requires, however, that the fixed-node constraint is relaxed during the projection, and therefore amounts to an expensive "nodal-release" approach. The approach has been successfully applied to the computation of low-lying excitations of bosonic systems [32, 93] and has also been used to compute tens of excited states of the fermionic, high-pressure liquid hydrogen in order to estimate its electrical conductivity [94].

If we move beyond a continuum formulation of QMC, the auxiliary-field quantum Monte Carlo (AFQMC) method by Zhang and coworkers [95] represents a very distinct, feasible alternative to DMC. In this approach, the random walk is in a space of single-particle Slater determinants, which are subject to a fluctuating external potential. The fermion-sign problem appears here in the form of a phase problem and is approximately eliminated by requiring that the phase of the determinant remains close to the phase of a trial wave function. The method is more expensive than DMC but has been applied to a variety of molecular and extended systems (mainly in the ground state) and appears to be less plagued by the phase constraint as compared to the effect of the fixed-node

approximation in DMC. A very recent review of AFQMC and its applications also to excited states can be found in Ref. [96].

Finally, we should mention another QMC approach in determinantal space, namely, the full configuration interaction quantum Monte Carlo method [97, 98] where a stochastic approach is used to select the important determinants in a full configuration interaction expansion. The method has been described in Chapter 6 together with its extension to excited states.

Bibliography

1 Foulkes, W.M.C., Mitas, L., Needs, R.J., and Rajagopal, G. (2001). Quantum Monte Carlo simulations of solids. *Rev. Mod. Phys.* 73: 33–83.
2 Lüchow, A. (2011). Quantum Monte Carlo methods. *WIREs Comput. Mol. Sci.* 1 (3): 388–402.
3 Austin, B.M., Zubarev, D.Y., and Lester, W.A. (2012). Quantum Monte Carlo and related approaches. *Chem. Rev.* 112: 263–288.
4 Hermann, J., Alfè, D., and Tkatchenko, A. (2017). Nanoscale stacked molecules are bound by collective charge fluctuations. *Nat. Commun.* 8: 14052.
5 Umrigar, C.J. and Filippi, C. (2005). Energy and variance optimization of many body wave functions. *J. Chem. Phys.* 94: 150201.
6 Umrigar, C.J., Toulouse, J., Filippi, C. et al. (2007). Alleviation of the fermion-sign problem by optimization of many-body wave functions. *Phys. Rev. Lett.* 98 (11): 110201.
7 Toulouse, J. and Umrigar, C.J. (2007). Optimization of quantum Monte Carlo wave functions by energy minimization. *J. Chem. Phys.* 126: 084102.
8 Zhao, L. and Neuscamman, E. (2017). A blocked linear method for optimizing large parameter sets in variational Monte Carlo. *J. Chem. Theory Comput.* 13 (6): 2604–2611.
9 Filippi, C., Zaccheddu, M., and Buda, F. (2009). Absorption spectrum of the green fluorescent protein chromophore: a difficult case for ab initio methods? *J. Chem. Theory Comput.* 5: 2074–2087.
10 Zhao, L. and Neuscamman, E. (2016). An efficient variational principle for the direct optimization of excited states. *J. Chem. Theory Comput.* 12: 3436–3440.
11 Shea, J.A.R. and Neuscamman, E. (2017). Size consistent excited states via algorithmic transformations between variational principles. *J. Chem. Theory Comput.* 13: 6078–6088.
12 Sorella, S. and Capriotti, L. (2010). Algorithmic differentiation and the calculation of forces by quantum Monte Carlo. *J. Chem. Phys.* 133 (23): 234111.
13 Filippi, C., Assaraf, R., and Moroni, S. (2016). Simple formalism for efficient derivatives and multi-determinant expansions in quantum Monte Carlo. *J. Chem. Phys.* 144 (19): 194105.
14 Assaraf, R., Moroni, S., and Filippi, C. (2017). Optimizing the energy with quantum Monte Carlo: a lower numerical scaling for Jastrow-Slater expansions. *J. Chem. Theory Comput.* 13: 5273–5281.
15 Dash, M., Moroni, S., Scemama, A., and Filippi, C. (2018). Perturbatively selected configuration-interaction wave functions for efficient geometry optimization in quantum Monte Carlo. *J. Chem. Theory Comput.* 14: 4176–4182.
16 Luo, Y., Zen, A., and Sorella, S. (2014). Ab initio molecular dynamics with noisy forces: validating the quantum Monte Carlo approach with benchmark calculations of molecular vibrational properties. *J. Chem. Phys.* 141: 194112.

17 Zen, A., Luo, Y., Mazzola, G. et al. (2015). Ab initio molecular dynamics simulation of liquid water by quantum Monte Carlo. *J. Chem. Phys.* 142: 144111.

18 Rios, P.L., Ma, A., Drummond, N. et al. (2006). Inhomogeneous backflow transformations in quantum Monte Carlo calculations. *Phys. Rev. E* 74: 066701.

19 Holzmann, M. and Moroni, S. (2019). Orbital-dependent backflow wave functions for real-space quantum Monte Carlo. *Phys. Rev. B* 99: 085121.

20 Casula, M. and Sorella, S. (2003). Geminal wave functions with Jastrow correlation: a first application to atom. *J. Chem. Phys.* 119: 6500.

21 Casula, M., Attaccalite, C., and Sorella, S. (2004). Correlated geminal wave function for molecules: an efficient resonating valence bond approach. *J. Chem. Phys.* 121: 7110.

22 Bajdich, M., Mitas, L., Wagner, L., and Schmidt, K. (2008). Pfaffian pairing and backflow wave functions. *Phys. Rev. B* 77: 115112.

23 Fracchia, F., Filippi, C., and Amovilli, C. (2012). Size-extensive wave functions for quantum Monte Carlo: a linear scaling generalized valence bond approach. *J. Chem. Theory Comput.* 8: 1943–1951.

24 Zulfikri, H., Amovilli, C., and Filippi, C. (2016). Multiple-resonance local wave functions for accurate excited states in quantum Monte Carlo. *J. Chem. Theory Comput.* 12 (3): 1157–1168.

25 Clay, R.C. and Morales, M.A. (2015). Influence of single particle orbital sets and configuration selection on multideterminant wave functions in quantum Monte Carlo. *J. Chem. Phys.* 142: 234103.

26 Per, M.C. and Cleland, D.M. (2017). Energy-based truncation of multi-determinant wave functions in quantum Monte Carlo. *J. Chem. Phys.* 146: 164101.

27 Robinson, P.J., Pineda Flores, S.D., and Neuscamman, E. (2017). Excitation variance matching with limited configuration interaction expansions in variational Monte Carlo. *J. Chem. Phys.* 147: 164114.

28 Kim, J., Baczewski, A.D., Beaudet, T.D. et al. (2018). QMCPACK: an open source ab initio quantum Monte Carlo package for the electronic structure of atoms, molecules and solids. *J. Phys. Condens. Matter* 30: 195901.

29 Scemama, A., Benali, A., Jacquemin, D. et al. (2018). Excitation energies from diffusion Monte Carlo using selected configuration interaction nodes. *J. Chem. Phys.* 149: 034108.

30 Pineda Flores, S.D. and Neuscamman, E. (2019). Excited state specific multi-Slater Jastrow wave functions. *Chem. A Eur. J.* 123 (8): 1487–1497.

31 Assaraf, R. and Caffarel, M. (1999). Zero-variance principle for Monte Carlo algorithms. *Phys. Rev. Lett.* 83: 4682.

32 Nightingale, M.P. and Melik-Alaverdian, V. (2001). Optimization of ground- and excited-state wave functions and van der Waals clusters. *Phys. Rev. Lett.* 87 (4): 043401.

33 Assaraf, R. and Caffarel, M. (2003). Zero-variance zero-bias principle for observables in quantum Monte Carlo: application to forces. *J. Chem. Phys.* 119: 10536.

34 Toulouse, J., Assaraf, R., and Umrigar, C.J. (2007). Zero-variance zero-bias quantum Monte Carlo estimators of the spherically and system-averaged pair density. *J. Chem. Phys.* 126: 244112.

35 Attaccalite, C. and Sorella, S. (2008). Stable liquid hydrogen at high pressure by a novel ab initio molecular-dynamics calculation. *Phys. Rev. Lett.* 100 (11): 114501.

36 Per, M.C., Snook, I.K., and Russo, S.P. (2012). Efficient calculation of unbiased expectation values in diffusion quantum Monte Carlo. *Phys. Rev. B* 86: 201107.

37 Burkatzki, M., Filippi, C., and Dolg, M. (2007). Energy-consistent pseudopotentials for quantum Monte Carlo calculations. *J. Chem. Phys.* 126 (23): 234105.

38 Burkatzki, M., Filippi, C., and Dolg, M. (2008). Energy-consistent small-core pseudopotentials for 3d-transition metals adapted to quantum Monte Carlo calculations. *J. Chem. Phys.* 129: 164115.

39 Trail, J.R. and Needs, R.J. (2017). Shape and energy consistent pseudopotentials for correlated electron systems. *J. Chem. Phys.* 146 (20): 204107.

40 Bennett, M.C., Melton, C.A., Annaberdiyev, A. et al. (2017). A new generation of effective core potentials for correlated calculations. *J. Chem. Phys.* 147: 224106.

41 Bennett, M.C., Wang, G., Annaberdiyev, A. et al. (2018). A new generation of effective core potentials from correlated calculations: 2nd row elements. *J. Chem. Phys.* 149: 104108.

42 Annaberdiyev, A., Wang, G., Melton, C.A. et al. (2018). A new generation of effective core potentials from correlated calculations: 3d transition metal series. *J. Chem. Phys.* 149: 134108.

43 Floris, F.M., Filippi, C., and Amovilli, C. (2012). A density functional and quantum Monte Carlo study of glutamic acid in vacuo and in a dielectric continuum medium. *J. Chem. Phys.* 137: 075102.

44 Floris, F.M., Filippi, C., and Amovilli, C. (2014). Electronic excitations in a dielectric continuum solvent with quantum Monte Carlo: acrolein in water. *J. Chem. Phys.* 140: 034109.

45 Daday, C., König, C., Neugebauer, J., and Filippi, C. (2014). Wave function in density functional theory embedding for excited states: which wave functions, which densities? *ChemPhysChem* 15 (15): 3205–3217.

46 Guareschi, R., Floris, F.M., Amovilli, C., and Filippi, C. (2014). Solvent effects on excited-state structures: a quantum Monte Carlo and density functional study. *J. Chem. Theory Comput.* 10: 5528–5537.

47 Guareschi, R., Zulfikri, H., Daday, C. et al. (2016). Introducing QMC/MMpol: quantum Monte Carlo in polarizable force fields for excited states. *J. Chem. Theory Comput.* 12: 1674–1683.

48 Doblhoff-Dier, K., Kroes, G.-J., and Libisch, F. (2018). Density functional embedding for periodic and nonperiodic diffusion Monte Carlo calculations. *Phys. Rev. B* 98: 085138.

49 Hammond, B.L., Lester, W.A., and Reynolds, P.J. (1994). *Monte Carlo Methods in Ab Initio Quantum Chemistry*. Singapore: WSPC.

50 Kolorenč, J. and Mitas, L. (2011). Applications of quantum Monte Carlo methods in condensed systems. *Rep. Prog. Phys.* 74 (2): 026502.

51 Busemeyer, B., Dagrada, M., Sorella, S. et al. (2016). Competing collinear magnetic structures in superconducting FeSe by first-principles quantum Monte Carlo calculations. *Phys. Rev. B* 94: 035108.

52 Chen, L. and Wagner, L.K. (2018). Quantum Monte Carlo study of the metal-to-insulator transition on a honeycomb lattice with 1/r interactions. *Phys. Rev. B* 97: 045101.

53 Dubecký, M., Mitas, L., and Jurečka, P. (2016). Noncovalent interactions by quantum Monte Carlo. *Chem. Rev.* 116: 5188–5215.

54 Toulouse, J., Assaraf, R., and Umrigar, C.J. (2016). Chapter Fifteen - Introduction to the variational and diffusion Monte Carlo methods. In: *Advances in Quantum Chemistry* (eds. P. E. Hoggan and T. Ozdogan), vol. 73. *Electron Correlation in Molecules - ab initio Beyond Gaussian Quantum Chemistry*, 285–314. Academic Press.

55 Ceperley, D.M. (1995). Path integrals in the theory of condensed helium. *Rev. Mod. Phys.* 67: 279–355.

56 Assaraf, R., Caffarel, M., and Scemama, A. (2007). Improved Monte Carlo estimators for the one-body density. *Phys. Rev. E* 75: 035701.

57 Per, M.C., Snook, I.K., and Russo, S.P. (2011). Zero-variance zero-bias quantum Monte Carlo estimators for the electron density at a nucleus. *J. Chem. Phys.* 135: 134112.

58 Umrigar, C.J., Nightingale, M.P., and Runge, K.J. (1993). A diffusion Monte Carlo algorithm with very small timestep errors. *J. Chem. Phys.* 99: 2865.

59 Rasch, K.M., Hu, S., and Mitas, L. (2014). Communication: Fixed-node errors in quantum Monte Carlo: interplay of electron density and node nonlinearities. *J. Chem. Phys.* 140: 041102.

60 Foulkes, W.M.C., Hood, R.Q., and Needs, R.J. (1999). Symmetry constraints and variational principles in diffusion quantum Monte Carlo calculations of excited-state energies. *Phys. Rev. B* 60: 4558–4570.

61 Schautz, F., Buda, F., and Filippi, C. (2004). Excitations in photoactive molecules from quantum Monte Carlo. *J. Chem. Phys.* 121: 5836–5844.

62 Prendergast, D., Nolan, M., Filippi, C. et al. (2001). Impact of electron-electron cusp on configuration interaction energies. *J. Chem. Phys.* 115 (4): 1626–1634.

63 Valsson, O. and Filippi, C. (2010). Photoisomerization of model retinal chromophores: insight from quantum Monte Carlo and multiconfigurational perturbation theory. *J. Chem. Theory Comput.* 6: 1275–1292.

64 Blunt, N.S. and Neuscamman, E. (2019). Excited-state diffusion Monte Carlo calculations: a simple and efficient two-determinant ansatz. *J. Chem. Theory Comput.* 15: 178–189.

65 Sorella, S. (2001). Generalized Lanczos algorithm for variational quantum Monte Carlo. *Phys. Rev. B* 64: 024512.

66 Casula, M., Attaccalite, C., and Sorella, S. (2004). Correlated geminal wave function for molecules: an efficient resonating valence bond approach. *J. Chem. Phys.* 121 (15): 7110–7126.

67 Schautz, F. and Filippi, C. (2004). Optimized Jastrow-Slater wave functions for ground and excited states: application to the lowest states of ethene. *J. Chem. Phys.* 120: 10931–10941.

68 Neuscamman, E., Umrigar, C.J., and Chan, G.K.-L. (2012). Optimizing large parameter sets in variational quantum Monte Carlo. *Phys. Rev. B* 85: 045103.

69 Toulouse, J. and Umrigar, C.J. (2008). Full optimization of Jastrow-Slater wave functions with application to the first-row atoms and homonuclear diatomic molecules. *J. Chem. Phys.* 128 (17): 174101.

70 Zimmerman, P.M., Toulouse, J., Zhang, Z. et al. (2009). Excited states of methylene from quantum Monte Carlo. *J. Chem. Phys.* 131: 124103.

71 Mussard, B., Coccia, E., Assaraf, R. et al. (2018). Time-dependent linear-response variational Monte Carlo. In: *Advances in Quantum Chemistry* (ed. P.E. Hoggan) vol. 76 of *Novel Electronic Structure Theory: General Innovations and Strongly Correlated Systems*, 255–270. Academic Press.

72 Umrigar, C.J., Wilson, K.G., and Wilkins, J.W. (1998). Optimized trial wave functions for quantum Monte Carlo calculations. *Phys. Rev. Lett.* 60: 1719–1722.

73 Grimes, R.M., Hammond, B.L., Reynolds, P.J., and Lester, W.A. (1986). Quantum Monte Carlo approach to electronically excited molecules. *J. Chem. Phys.* 85: 4749–4750.

74 Williamson, A.J., Grossman, J.C., Hood, R.Q. et al. (2002). Quantum Monte Carlo calculations of nanostructure optical gaps: application to silicon quantum dots. *Phys. Rev. Lett.* 89: 196803.

75 Hood, R.Q., Kent, P.R.C., Needs, R.J., and Briddon, P.R. (2003). Quantum Monte Carlo study of the optical and diffusive properties of the vacancy defect in diamond. *Phys. Rev. Lett.* 91: 076403.

76 Puzder, A., Williamson, A.J., Grossman, J.C., and Galli, G. (2003). Computational studies of the optical emission of silicon nanocrystals. *J. Am. Chem. Soc.* 125: 2786–2791.

77 El Akramine, O., Kollias, A.C., and Lester, W.A. (2003). Quantum Monte Carlo study of singlet-triplet transition in ethylene. *J. Chem. Phys.* 119: 1483–1488.

78 Aspuru-Guzik, A., El Akramine, O., Grossman, J.C., and Lester, W.A. (2004). Quantum Monte Carlo for electronic excitations of free-base porphyrin. *J. Chem. Phys.* 120: 3049–3050.

79 Drummond, N.D., Williamson, A.J., Needs, R.J., and Galli, G. (2005). Electron emission from diamondoids: a diffusion quantum Monte Carlo study. *Phys. Rev. Lett.* 95: 096801.

80 Bande, A., Lüchow, A., Della Sala, F., and Görling, A. (2006). Rydberg states with quantum Monte Carlo. *J. Chem. Phys.* 124: 114114.

81 Vincent, J.E., Kim, J., and Martin, R.M. (2007). Quantum Monte Carlo calculations of the optical gaps of Ge nanoclusters using core-polarization potentials. *Phys. Rev. B* 75: 045302.

82 Tiago, M.L., Kent, P.R.C., Hood, R.Q., and Reboredo, F.A. (2008). Neutral and charged excitations in carbon fullerenes from first-principles many-body theories. *J. Chem. Phys.* 129: 084311.

83 Marsusi, F., Sabbaghzadeh, J., and Drummond, N.D. (2011). Comparison of quantum Monte Carlo with time-dependent and static density-functional theory calculations of diamondoid excitation energies and Stokes shifts. *Phys. Rev. B* 84: 245315.

84 Send, R., Valsson, O., and Filippi, C. (2011). Electronic excitations of simple cyanine dyes: reconciling density functional and wave function methods. *J. Chem. Theory Comput.* 7: 444–455.

85 Robinson, P.J., Pineda Flores, S.D., and Neuscamman, E. (2017). Excitation variance matching with limited configuration interaction expansions in variational Monte Carlo. *J. Chem. Phys.* 147 (16): 164114.

86 Cordova, F., Doriol, L.J., Ipatov, A. et al. (2007). Troubleshooting time-dependent density-functional theory for photochemical applications: oxirane. *J. Chem. Phys.* 127: 164111.

87 Tapavicza, E., Tavernelli, I., Rothlisberger, U. et al. (2008). Mixed time-dependent density-functional theory/classical trajectory surface hopping study of oxirane photochemistry. *J. Chem. Phys.* 129: 124108.

88 Dubecký, M., Derian, R., Mitas, L., and Stich, I. (2010). Ground and excited electronic states of azobenzene: a quantum Monte Carlo study. *J. Chem. Phys.* 133: 244301.

89 Guareschi, R. and Filippi, C. (2013). Ground- and excited-state geometry optimization of small organic molecules with quantum Monte Carlo. *J. Chem. Theory Comput.* 9 (12): 5513–5525.

90 Amovilli, C., Filippi, C., and Floris, F.M. (2008). Quantum Monte Carlo formulation of volume polarization in dielectric continuum theory. *J. Chem. Phys.* 129 (24): 244106.

91 Filippi, C., Buda, F., Guidoni, L., and Sinicropi, A. (2012). Bathochromic shift in green fluorescent protein: a puzzle for QM/MM approaches. *J. Chem. Theory Comput.* 8: 112–124.

92 Valsson, O., Campomanes, P., Tavernelli, I. et al. (2013). Rhodopsin absorption from first principles: Bypassing common pitfalls. *J. Chem. Theory Comput.* 9: 2441–2454.

93 Ceperley, D.M. and Bernu, B. (1988). The calculation of excited state properties with quantum Monte Carlo. *J. Chem. Phys.* 89: 6316.

94 Lin, F., Morales, M.A., Delaney, K.T. et al. (2009). Electrical conductivity of high-pressure liquid hydrogen by quantum Monte Carlo methods. *Phys. Rev. Lett.* 103: 256401.

95 Zhang, S. and Krakauer, H. (2003). Quantum Monte Carlo method using phase-free random walks with Slater determinants. *Phys. Rev. Lett.* 90: 136401.

96 Zhang, S. (2018). Ab initio electronic structure calculations by auxiliary-field quantum Monte Carlo. In: *Handbook of Materials Modeling* (eds. W. Andreoni and S. Yip), 1–27. Cham: Springer.

97 Booth, G.H., Thom, A.J.W., and Alavi, A. (2009). Fermion Monte Carlo without fixed nodes: a game of life, death, and annihilation in Slater determinant space. *J. Chem. Phys.* 131 (5): 054106.

98 Cleland, D., Booth, G.H., and Alavi, A. (2010). Communications: Survival of the fittest: accelerating convergence in full configuration-interaction quantum Monte Carlo. *J. Chem. Phys.* 132 (4): 041103.

9

Multi-Reference Configuration Interaction

Felix Plasser[1] *and Hans Lischka*[2]

[1] *Department of Chemistry, Loughborough University, Loughborough, LE11 3TU, United Kingdom*
[2] *Department of Chemistry and Biochemistry, Texas Tech University, Memorial Circle & Boston, Lubbock, TX 79409, USA*

Abstract

The multireference configuration interaction (MRCI) method is a simple and robust way of treating non-dynamic and dynamic correlation on an equal footing. In its original form, it is based on a straightforward linear expansion of the wave function into Slater determinants or configuration state functions (CSFs) and the application of the variational principle. In this chapter the construction of this expansion space by means of reference configurations and a hierarchy of excitations thereof is discussed and a classification of the corresponding orbital spaces is provided. The violation of size-extensivity is explained using a simple model example and the possibilities of *a posteriori* and intrinsic corrections to restore approximate size-extensivity are presented. Contraction schemes based on internally contracted (ic)-MRCI, which alleviate the high computational demand of the uncontracted MRCI, are explained. The description of configuration selection schemes and different options for choosing orbitals to be used in the MRCI conclude this chapter.

9.1 Introduction

The multi-reference configuration interaction (MRCI) method is a way of treating non-dynamic and dynamic electron correlation on an even footing. MRCI operates by initially constructing a space of reference configurations, exciting electrons out of these configurations, and solving the electronic Schrödinger equation by variationally minimizing the energy in the resulting configuration space. In its uncontracted (uc) variant, excitations are performed with respect to each individual reference configuration. In addition, different contraction schemes have been developed to cope with the rapidly increasing computational cost of uc-MRCI through a reduction of the number of configurations, from which the most popular one is the internal contraction (ic). The MRCI method is conceptually simple and, in its uncontracted form, the working equations are equivalent to the single reference (SR) case with the exception that a larger set of configurations is considered. As a consequence, the MRCI method is very flexible, both in the choice of configuration space and the construction of the molecular orbitals and can be applied to a wide variety of molecular systems. Furthermore, MRCI is a variational method, which means that the energy obtained is never lower than the true ground state energy. The variationality of MRCI also leads to the fact that wave function properties, such as dipole moments, and energy gradients can be

Quantum Chemistry and Dynamics of Excited States: Methods and Applications,
First Edition. Edited by Leticia González and Roland Lindh.

computed fairly easily, particularly in the case of uc-MRCI. The main downside of MRCI is its lack of size-extensivity whenever the CI expansion is truncated at a specific excitation level. Practically this means that the description deteriorates as the size of the molecule increases.

This chapter is intended as a compact and easy-to-read introduction to the topic of MRCI. For a more comprehensive discussion, we refer the reader to the literature [1–5]. Several of the mathematical foundations and procedures are being discussed in other chapters of this book and we will refer to them as needed. Here, we start by discussing the basic definitions and concepts, focusing on uc-MRCI (Section 9.2). We continue by specifying different variants of MRCI in Section 9.3 and finish by listing popular implementations (Section 9.4).

9.2 Basics

In this section, we first discuss the basic points underlying MRCI starting with the main working equations of CI and showing how these lead to a method satisfying the variational principle (Section 9.2.1). In Section 9.2.2, we discuss how truncated CI invariably leads to a lack of size-extensivity. We continue by explaining the construction of configuration spaces (Section 9.2.3) and the different choices for the underlying many-electron basis functions (Section 9.2.4). As a final point, we outline the workflow in practical uc-MRCI computations considering the determination of the energy and its gradient with respect to variations in the geometry (Section 9.2.5).

9.2.1 Configuration Interaction and the Variational Principle

The general idea of the configuration interaction (CI) method is to compute the many-electron wave function of an electronic state as a linear combination of pre-defined many-electron basis functions. Explicitly, the wave function $|\Psi^\alpha\rangle$ of state α is written as

$$|\Psi^\alpha\rangle = \sum_{k=1}^{N_{\mathrm{CI}}} c_k^\alpha |\Phi_k\rangle \tag{9.1}$$

Here, the $|\Phi_k\rangle$ are the many-electron basis functions representing the electronic configurations, which will be discussed in more detail in Section 9.2.4. The c_k^α are the CI-coefficients, which are usually collected in the CI-vector $\mathbf{c}^\alpha$ whose length is given by N_{CI}. To derive the main working equation of CI, we start with the electronic Schrödinger equation

$$\hat{H}|\Psi^\alpha\rangle = E^\alpha|\Psi^\alpha\rangle \tag{9.2}$$

where $\hat{H}$ is the electronic Hamiltonian operator and E^α is the energy of state α. After insertion of Eq. (9.1) one obtains

$$\sum_{k=1}^{N_{\mathrm{CI}}} c_k^\alpha \hat{H}|\Phi_k\rangle = \sum_{k=1}^{N_{\mathrm{CI}}} c_k^\alpha E^\alpha |\Phi_k\rangle. \tag{9.3}$$

In the next step, we project the above equation onto $\langle\Phi_l|$ – in other words, we left-multiply the equation by the function Φ_l and integrate over all space – to obtain

$$\sum_{k=1}^{N_{\mathrm{CI}}} c_k^\alpha \langle\Phi_l|\hat{H}|\Phi_k\rangle = \sum_{k=1}^{N_{\mathrm{CI}}} c_k^\alpha E^\alpha \langle\Phi_l|\Phi_k\rangle. \tag{9.4}$$

To proceed, we define the CI-matrix $\mathbf{H}$ whose elements are defined as

$$H_{lk} = \langle\Phi_l|\hat{H}|\Phi_k\rangle. \tag{9.5}$$

In addition, we assume that the chosen many-electron basis functions form an orthonormal set so that $\langle \Phi_l | \Phi_k \rangle = \delta_{lk}$. Inserting these two expressions into Eq. (9.4) leads to

$$\sum_{k=1}^{N_{\mathrm{CI}}} H_{lk} c_k^{\alpha} = E^{\alpha} c_l^{\alpha}. \tag{9.6}$$

This is a set of N_{CI} different equations, one for each possible index of l. This set of equations can be conveniently written in matrix form to yield

$$\mathbf{H}\mathbf{c}^{\alpha} = E^{\alpha}\mathbf{c}^{\alpha}. \tag{9.7}$$

This is the main working equation of CI. The significance of this is that the initial operator eigenvalue equation (9.2) has been turned into a matrix eigenvalue equation. Equation (9.7) can be addressed by standard linear algebra techniques whereas the direct solution of the many-electron Schrödinger equation (9.2) is a daunting task that can be tackled only in very specific cases.

Equation (9.7) holds equivalently for the SR and MR cases where the only difference is that different sets of configurations are considered. Furthermore, there is no specific requirement on the underlying orbitals used. Equation (9.7) also provides a direct route to the computation of excited states in the uc-MRCI method. Whereas the ground state corresponds to the lowest energy eigenvalue of $\mathbf{H}$, excited states simply correspond to higher energy eigenvalues. These properties reflect the conceptual simplicity of the MRCI method and contribute to its flexibility.

In practice, the dimension of the CI vector N_{CI} may well exceed 10^9 configurations. In this case, it is clearly not possible to store the whole CI matrix $\mathbf{H}$ in memory or even on disk. For this purpose, the CI matrix is generated on-the-fly from the MO integrals in the direct CI procedure [6].

One of the main properties of the CI method is the fact that it is variational. It fulfills two important properties in this context. First, if we construct any trial wave function according to Eq. (9.1)

$$|\Psi^{\mathrm{trial}}\rangle = \sum_{k=1}^{N_{\mathrm{CI}}} c_k^{\mathrm{trial}} |\Phi_k\rangle \tag{9.8}$$

then its energy expectation value

$$E^{\mathrm{trial}} = \frac{\langle \Psi^{\mathrm{trial}} | \hat{H} | \Psi^{\mathrm{trial}} \rangle}{\langle \Psi^{\mathrm{trial}} | \Psi^{\mathrm{trial}} \rangle} \geq E^{0,\mathrm{exact}} \tag{9.9}$$

will always be greater or equal to the true ground state energy of the system $E^{0,\mathrm{exact}}$. This simply follows from the fact that the CI energy is computed as an expectation value. Furthermore, the energy of any trial wave function is always greater or equal to the lowest eigenvalue E^0 of the CI-matrix $\mathbf{H}$, which corresponds to the CI energy of the ground state. This can be readily seen if the expectation value is given in matrix form

$$E^{\mathrm{trial}} = \frac{(\mathbf{c}^{\mathrm{trial}})^{\mathrm{T}} \mathbf{H} \mathbf{c}^{\mathrm{trial}}}{(\mathbf{c}^{\mathrm{trial}})^{\mathrm{T}} \mathbf{c}^{\mathrm{trial}}} \geq E^0. \tag{9.10}$$

In the case of a non-degenerate ground state, the equality $E^{\mathrm{trial}} = E^0$ is obtained only if $|\Psi^{\mathrm{trial}}\rangle$ corresponds to the ground-state wave function. In the case, of a degenerate ground state, this relation holds whenever $|\Psi^{\mathrm{trial}}\rangle$ is formed as a linear combination of the degenerate ground-state wave functions. In summary, it holds that the CI ground-state energy is the lowest possible energy expectation value for any possible trial wave function within the configuration space and that this energy will always be greater or equal to the exact ground state energy. Note that neither of these statements holds for the coupled-cluster method.

The variational theorem can be generalized to excited states [2]. To see this, we start with the CI-matrix $\mathbf{H}$ whose lowest eigenvalues are $E^0, E^1, E^2, \ldots$ as indicated in Eq. (9.7). Next, a basis function is added to the configuration space and a new Hamiltonian matrix with an extra row and an extra column is generated

$$\tilde{\mathbf{H}} = \begin{pmatrix} \mathbf{H} & \tilde{\mathbf{v}} \\ \tilde{\mathbf{v}}^{\mathrm{T}} & \tilde{\epsilon} \end{pmatrix} \tag{9.11}$$

The Hylleraas–Undheim theorem [7] now states that for the eigenvalues of $\tilde{\mathbf{H}}$, denoted $\tilde{E}^0, \tilde{E}^1, \tilde{E}^2, \ldots$, the following relation holds

$$\tilde{E}^0 \leq E^0 \leq \tilde{E}^1 \leq E^1 \leq \tilde{E}^2 \leq E^2 \leq \ldots \tag{9.12}$$

This can be understood in the following way

- If $\tilde{\epsilon}$ is lower than a specific eigenvalue E^α, it will directly contribute to a new lower eigenvalue of $\tilde{\mathbf{H}}$.
- If $\tilde{\epsilon} \geq E^\alpha$ it will either leave the eigenvalue unchanged or lower it through off-diagonal coupling terms.

The first relation of Eq. (9.12), that is $\tilde{E}^0 \leq E^0$, corresponds to the usual variational theorem: the lowest eigenvalue either stays the same or becomes lower as the variational space is increased. Equation (9.12) shows that a similar relation holds for every individual eigenvalue. The energy of any specific state E^α can never increase as the variational space is increased. This argument can be extended and it can be understood that the CI eigenvalues are upper bounds to the exact eigenvalues (the solutions of Eq. (9.2)) and it holds that

$$E^0 \geq E^{0,\mathrm{exact}} \tag{9.13}$$

$$E^1 \geq E^{1,\mathrm{exact}} \tag{9.14}$$

$$E^2 \geq E^{2,\mathrm{exact}} \tag{9.15}$$

$$\ldots$$

Finally, two words of caution are in order. First, Eqs. (9.13)–(9.15) only hold for the total energies of the states and it is not possible to make a similar statement about excitation energies. Second, these equations only apply if all excited states are obtained as eigenvalues of the same matrix.

An alternative way of deriving the CI working equations (9.7) starts with Eq. (9.10) and minimizes E^{trial} explicitly invoking the variational principle. For this purpose, a Lagrange multiplier formalism can be used to make E^{trial} stationary with respect to variations in the CI coefficients and enforce a normalized wave function [8]. The fact that the CI energy is stationary with respect to the CI coefficients facilitates the computation of wave function properties, gradients and non-adiabatic couplings [5, 8]. This holds also in the MR case and implementations of gradients and non-adiabatic couplings have been reported for uncontracted MRCI [9–11] and are widely used.

9.2.2 The Size-Extensivity Problem of Truncated CI

In practical cases, it is most common to truncate the CI expansion at a specific excitation level. In the SR case one might use singles (CIS), doubles (CISD), triples (CISDT) while in the MR case, the MRCIS and MRCISD methods are commonly used. The main practical problem of these truncated methods is their lack of size-extensivity. Here, the term "extensive" comes from thermodynamics where the term is used to describe a physical quantity whose magnitude is proportional to the size

of the system. The electronic correlation energy of a system is clearly supposed to be an extensive property of the system. A computational method is called "size-extensive" if its correlation energy scales correctly with the size of the system [12, 13], i.e., the correlation energy is truly an extensive property. A different way of viewing size-extensivity is the requirement that the correlation energy per particle does not tend to zero as the system size increases [14]. For practical purposes, it is often easier to think about the related concept of size-consistency. Here, a size-consistent method is a method where a computation of non-interacting systems will lead to the same energy as the sum of calculations on the isolated systems [15].

Before proceeding it is worth noting that two commonly used CI methods, SRCIS (if based on canonical orbitals) and full CI, are indeed size-extensive. The Brillouin theorem states that singly excited configurations do not interact with the converged Hartree–Fock wave function. As a consequence, the SRCIS ground state is equivalent to the Hartree–Fock ground state, which is size-extensive [16, 17]. The size-extensivity of the SRCIS excited state energies follows, for example, from the fact that these are equivalent to the energies of the size extensive coupled cluster singles (CCS) method. Full CI is also size-extensive but this happens at the cost of exponential scaling with the system size.

To explain the lack of size consistency of truncated CI, we may start by the following qualitative argument:

- CISD is equivalent to full CI for a two-electron system.
- CISD is no longer equivalent to full CI for two non-interacting two-electron systems.

Thus, the description of two non-interacting systems is different from the description of the individual systems, which violates the definition of size-consistency given above.

To understand the problem in some more detail, it is convenient to start with a hydrogen molecule in a minimal basis set or any other two-electron system, see also Ref. [13]. The hydrogen molecule possesses two molecular orbitals, the bonding σ_g orbital and the anti-bonding σ_u^* orbital. The Hartree–Fock reference determinant Φ_0 corresponds to the situation where the σ_g orbital is doubly occupied. For symmetry reasons this determinant can only interact with the doubly excited configuration Φ_D. The wave function for an isolated hydrogen atom is thus given as:

$$|\Psi\rangle = c_0|\Phi_0\rangle + c_D|\Phi_D\rangle. \tag{9.16}$$

It is a linear combination of the reference determinant and a doubly excited configuration. This wave function can be completely described by the CISD method. CISD is equivalent to full CI for this system and this holds for any two-electron system independent of the basis set.

If two non-interacting hydrogen molecules are present, one can – ignoring the Pauli principle – write the total wave function Ψ^{AB} of the system as a product of the wave functions of the individual hydrogen atoms Ψ^A and Ψ^B

$$|\Psi^{AB}\rangle = |\Psi^A\rangle|\Psi^B\rangle = \tag{9.17}$$

$$(c_0|\Phi_0^A\rangle + c_D|\Phi_D^A\rangle)(c_0|\Phi_0^B\rangle + c_D|\Phi_D^B\rangle) = \tag{9.18}$$

$$c_0{}^2|\Phi_0^A\rangle|\Phi_0^B\rangle + c_0c_D(|\Phi_0^A\rangle|\Phi_D^B\rangle + |\Phi_D^A\rangle|\Phi_0^B\rangle) + c_D{}^2|\Phi_D^A\rangle|\Phi_D^B\rangle \tag{9.19}$$

The first term in Eq. (9.19) is the closed-shell reference of the combined system. The two following configurations are doubly excited with respect to this reference and the final configuration shown is quadruply excited. In a short-hand notation, we can rewrite Eq. (9.19) as

$$|\Psi^{AB}\rangle = c_0{}^2\left[|\Phi_0^{AB}\rangle + \frac{c_D}{c_0}\sqrt{2}|\Phi_D^{AB}\rangle + \frac{c_D{}^2}{c_0{}^2}|\Phi_Q^{AB}\rangle\right] \tag{9.20}$$

Here, we have constructed normalized functions for the combined system representing the reference, doubly and quadruply excited configurations, respectively, using the definitions

$$|\Phi_0^{AB}\rangle = |\Phi_0^A\rangle|\Phi_0^B\rangle \tag{9.21}$$

$$|\Phi_D^{AB}\rangle = 1/\sqrt{2}(|\Phi_0^A\rangle|\Phi_D^B\rangle + |\Phi_D^A\rangle|\Phi_0^B\rangle) \tag{9.22}$$

$$|\Phi_Q^{AB}\rangle = |\Phi_D^A\rangle|\Phi_D^B\rangle. \tag{9.23}$$

Quadruply excited configurations naturally arise even though the individual hydrogen molecules are only doubly excited. It is no longer possible to construct a wave function of the form of Eq. (9.19) using the CISD method, as only the first two terms can be included. This means that the CISD wave function is not size-consistent as the description of the combined system is different from the description of the individual systems.

In the case of only two hydrogen molecules in a minimal basis set, there is only one possible quadruple excitation and we can assume that its weight is small considering that $c_0 \gg c_D$. However, the situation changes once the number of particles is increased for combinatorial reasons. If there are N hydrogen molecules, then there are N possibilities of exciting one of them and $\frac{N(N-1)}{2}$ possibilities of exciting two of them at the same time. As a consequence, for N hydrogen molecules Eq. (9.19) takes the form

$$|\Psi^{AB\ldots N}\rangle = {c_0}^N \left[|\Phi_0^{AB\ldots N}\rangle + \frac{c_D}{c_0}\sqrt{N}|\Phi_D^{AB\ldots N}\rangle + \frac{{c_D}^2}{{c_0}^2}\sqrt{\frac{N(N-1)}{2}}|\Phi_Q^{AB\ldots N}\rangle + \ldots \right]. \tag{9.24}$$

Here, $|\Phi_0^{AB\ldots N}\rangle$ refers to the product of the N reference determinants in analogy to Eq. (9.21) while $|\Phi_D^{AB\ldots N}\rangle$ and $|\Phi_Q^{AB\ldots N}\rangle$ are normalized linear combinations of all doubly excited and pairs of doubly excited configurations, respectively. Eq. (9.24) shows that, no matter how small C_D is, the contribution of the quadruples will become important when N is increased. More generally, as the system size increases, higher excitations will become dominant in the wave function expansion and the weight of the reference determinant and the lower excitations ultimately tends toward zero.

The higher excitation terms in Eq. (9.24) are often denoted "disconnected" excitations, using the language of diagrammatic perturbation theory [13]. The coefficients of these excitations are not determined by independent wave function parameters but their magnitude depends on the CI-coefficient c_D of the doubly excited configuration of the isolated system. This observation is at the heart of coupled-cluster theory: it is enough to know the coefficients of the double excitations to compute higher excitations as products of the lower excitations. In the case of MRCI, it is possible to use similar arguments for approximate extensivity corrections, see Section 9.3.2.

9.2.3 Multi-Reference Configuration Spaces

MRCI computations are usually carried out by first constructing a space of reference configurations and subsequently allowing a fixed number of excitations out of this reference space. The reference space is intended to cover non-dynamic electron correlation whereas the excitations out of it represent dynamic correlation. In this section, we will first discuss the separation of the molecular orbitals (MOs) into different spaces and will subsequently show how the different electronic configurations are constructed using these orbital spaces.

Before setting up an MRCI computation, it is necessary to divide the MOs into different subspaces as shown in Figure 9.1. At the bottom of Figure 9.1 are the frozen core orbitals. These remain doubly occupied in the reference space as well as the actual MRCI configuration space.

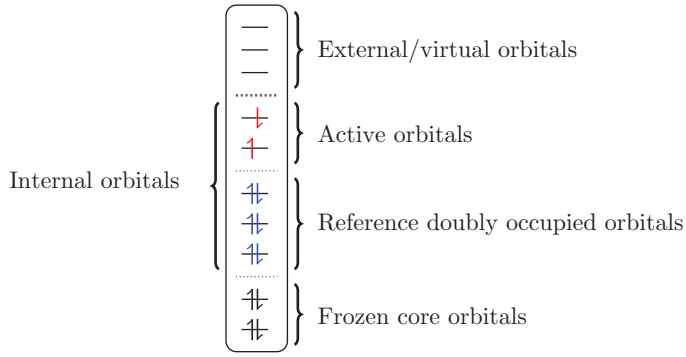


Figure 9.1 Division of the MRCI molecular orbital space into different subspaces.

The default is to consider the inner electron shells of the system, e.g., the 1s orbitals of second row elements, as frozen core orbitals. But sometimes more orbitals are frozen to save computational cost. For example, it is possible to freeze all σ electrons in a π-system [18]. Next, the reference doubly occupied (refdocc) orbitals follow. These are doubly occupied in all reference configurations but excitations out of these orbitals are allowed when generating the MRCI wave function. The next set of orbitals considered are the active orbitals. These possess various occupation patterns already in the reference configurations. It is also common to refer to the combined set of active and reference doubly occupied orbitals as internal orbitals as indicated in Figure 9.1 on the left. The last space to be considered are the external or virtual orbitals. These are unoccupied in the reference configurations but are populated by one or two electrons during the MRCI procedure. Finally, it is also possible to freeze and, thus, effectively delete virtual orbitals from the computation. Frozen virtual orbitals do not affect the MRCI results at all and are, therefore, not shown in Figure 9.1.

Having defined the different orbital spaces, we can now proceed by discussing the MRCI configuration space, as shown in Figure 9.2. At the heart of the MRCI computation is the reference space. The reference space is constructed by exciting electrons within the active orbital space. It is common to construct the reference space as a complete active space, i.e., by allowing all possible occupations within the active orbital space, in analogy to the CASSCF wave function (see Chapter 6). Alternatively, additional occupation restrictions, as discussed also in Chapter 6, can be imposed to reduce the size of the reference space and ultimately the configuration space. The reference space is shown in the upper left of Figure 9.2 representing a CAS(2,2) space.

Once the reference space is defined, the MRCI wave function is defined by exciting electrons out of this space. In MRCI there are four distinct types of excitations:

- refdocc → active
- active → external
- refdocc → external
- active → active

In the following, we will consider the uc-MRCISD method where excitations are performed with respect to the individual reference configurations and in total two excitations are allowed. This, first, leads to various singly excited configurations, which are shown on the upper right of Figure 9.2. The first two are of refdocc → active type, the following is active → external, and finally three refdocc → external configurations are shown. The fourth type of excitation (active → active) only comes into play for incomplete reference spaces and is not shown here. In the case of doubly excited configurations, the situation is even more complicated as any combination of

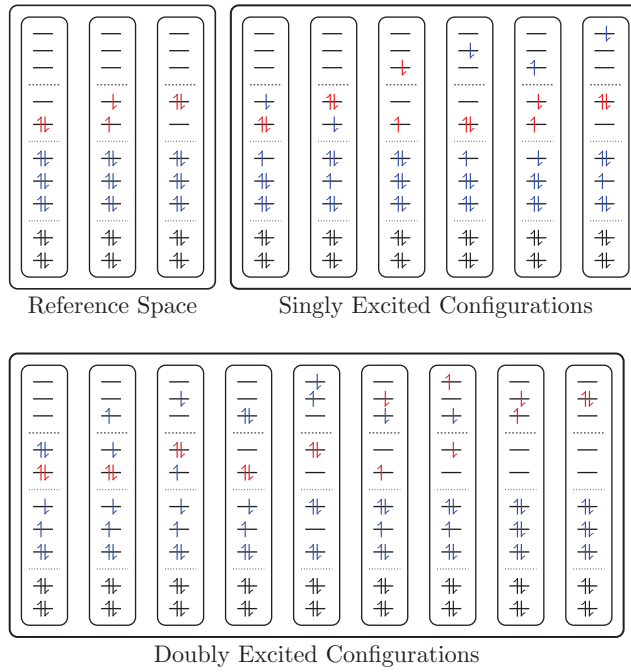


Figure 9.2 Division of the uc-MRCI configuration space into a reference space and the spaces of singly and doubly excited configurations. Active electrons are shown in red; electrons originating from the reference doubly occupied orbitals are shown in blue.

the above-mentioned excitation types is possible. In the bottom panel of Figure 9.2, a number of different doubly excited configurations are shown where different numbers of refdocc and active electrons are distributed over the different spaces.

Considering the representation of Figure 9.2, the MRCISD wave function can be written in the form

$$|\Psi\rangle = \sum_{k=1}^{N_{\text{ref}}} c_k^{\text{ref}} |\Phi_k^{\text{ref}}\rangle + \sum_l c_l^{\text{S}} |\Phi_l^{\text{S}}\rangle + \sum_m c_m^{\text{D}} |\Phi_m^{\text{D}}\rangle \tag{9.25}$$

where $|\Phi_k^{\text{ref}}\rangle$ are the reference configurations and $|\Phi_l^{\text{S}}\rangle$ and $|\Phi_m^{\text{D}}\rangle$ are the singly and doubly excited configurations, respectively.

Figure 9.2 illustrates the complexity of the MRCI wave function. Whereas in the single-reference case only one type of excitation exists (from occupied to virtual orbitals), three different orbital spaces and excitations between them have to be considered for MRCI. A practical problem in this context is that the same final excited configuration can often be reached through excitations from different references. For example, the first singly excited configuration in Figure 9.2 could be constructed in three different ways:

- starting from the first reference configuration and exciting an electron from the refdocc orbital to the higher-lying active orbital (as indicated by the colors in Figure 9.2),
- starting from the second reference configuration and exciting an electron from the refdocc orbital to the lower-lying active orbital,
- or starting from the third reference configuration and exciting two electrons (refdocc → active and active → active).

Thus, depending on the choice made, this configuration would either be counted as a $|\Phi_l^S\rangle$ or a $|\Phi_m^D\rangle$ configuration and it is clear that the expansion (9.25) creates linear dependencies. In the case of MRCI this is mainly an "accounting problem" and different methods have been developed for removing redundancies in the expansion space and creating a unique list of configurations [19, 20]. By contrast, the existence of different orbital spaces is a major hindrance in the development of MR coupled cluster approaches as no uniquely defined "Fermi vacuum" exists that can be used as a starting point for the coupled cluster expansion [5, 14].

The division into reference, singly and doubly excited configurations, as shown in Figure 9.2, is most convenient when constructing the configuration space. As opposed to this representation, it is also common to divide the configuration space according to the number of electrons in internal orbitals, see Figure 9.3. In this representation one starts with the internal space, which is composed of all configurations where the electrons are in internal (refdocc and active) orbitals. These correspond to the reference configurations and all configurations obtained through refdocc → active excitations shown on the left and right, respectively, in the upper panel of Figure 9.3. In the case of an incomplete reference space, also active → active excitations are included here. The 1-external space corresponds to all configurations where one external orbital is singly occupied (Figure 9.3, lower left). These configurations are either obtained by single excitations out of the internal space or double excitations where one electron stays in the internal space and the other one is excited out of it. The 2-external space is obtained by double excitations. It is usually the largest part of the configuration space. Finally, it is possible to differentiate whether the two electrons excited into the

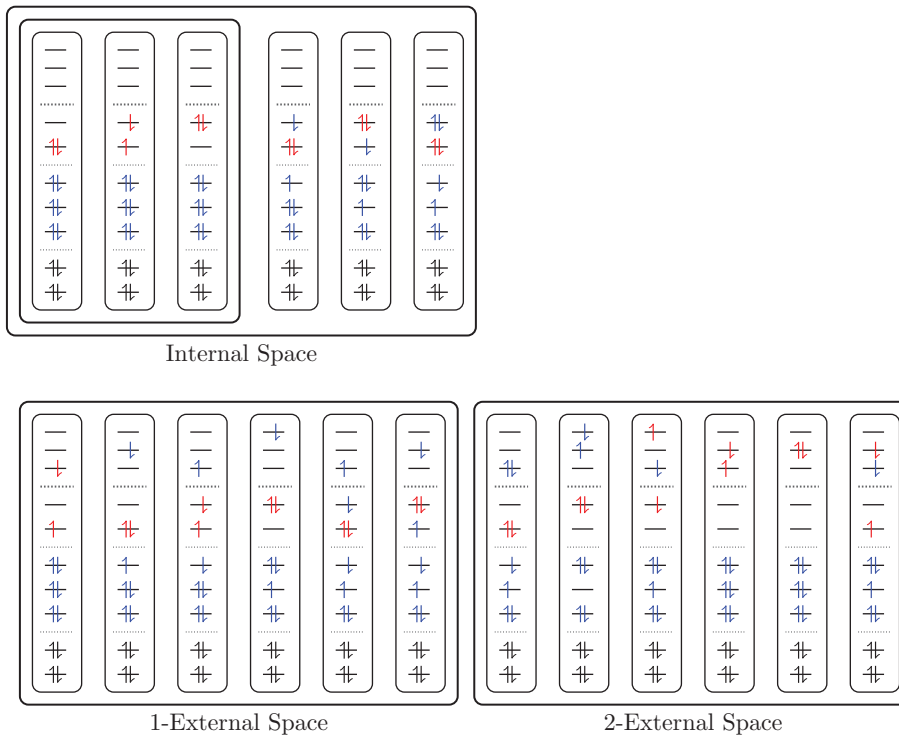


Figure 9.3 Division of the uc-MRCI configuration space into an internal space and configurations with one or two electrons in external orbitals. Active electrons are shown in red; electrons originating from reference doubly occupied orbitals are shown in blue.

external space are of the same or opposite spin. In the example shown, the two electrons have the same spin for all except the last 2-external configuration.

In light of the representation of Figure 9.3, the MRCISD wave function can be written in the form

$$|\Psi\rangle = \sum_k c_k^{\text{int}} |\Phi_k^{\text{int}}\rangle + \sum_l c_l^{\text{1-ext}} |\Phi_l^{\text{1-ext}}\rangle + \sum_m c_m^{\text{2-ext}} |\Phi_m^{\text{2-ext}}\rangle \tag{9.26}$$

where $|\Phi_k^{\text{int}}\rangle$ are the internal configurations and $|\Phi_l^{\text{1-ext}}\rangle$ and $|\Phi_m^{\text{2-ext}}\rangle$ are the configurations with one and two electrons in external orbitals, respectively. As opposed to Eq. (9.25), this is a unique representation where every configuration fits only into one of the categories. A more compact description is given by only separating the internal and external configurations

$$|\Psi\rangle = \sum_k c_k^{\text{int}} |\Phi_k^{\text{int}}\rangle + \sum_l c_l^{\text{ext}} |\Phi_l^{\text{ext}}\rangle, \tag{9.27}$$

and this separation is used in the definition of some extensivity corrected methods as discussed in Section 9.3.2.

9.2.4 Many-Electron Basis Functions: Determinants and CSFs

In the above section, we have discussed different electron configurations without discussing the nature of the underlying many-electron basis functions. This is the focus of this section. In practice, there are two common options: Slater determinants and configuration state functions (CSFs). The difference between them is that Slater determinants are only eigenfunctions of the $\hat{S}_z$ operator, i.e., the spin-operator projected onto one spatial dimension, while CSFs are also eigenfunctions of the total spin operator $\hat{S}^2$. Slater determinants are uncorrelated aside from Fermi correlation mandated by the Pauli principle. CSFs encode additional non-trivial static correlation effects.

The simplest way of constructing a configuration for CI is in the form of a Slater determinant

$$\Phi^{\text{S}}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) = \frac{1}{\sqrt{n!}} \begin{vmatrix} \phi_1(\mathbf{x}_1) & \phi_2(\mathbf{x}_1) & \dots & \phi_n(\mathbf{x}_1) \\ \phi_1(\mathbf{x}_2) & \phi_2(\mathbf{x}_2) & \dots & \phi_n(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(\mathbf{x}_n) & \phi_2(\mathbf{x}_n) & \dots & \phi_n(\mathbf{x}_n) \end{vmatrix} \tag{9.28}$$

where the ϕ_i are the molecular orbitals, $\mathbf{x}_i$ refers to the coordinates of the ith electron, and n is the number of electrons. The defining feature of a Slater determinant is that it yields an antisymmetric wave function compliant with the Pauli principle, i.e., $\Phi^{\text{S}}(\mathbf{x}_2, \mathbf{x}_1, \dots, \mathbf{x}_n) = -\Phi^{\text{S}}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$. This follows from the fact that the exchange of two electrons corresponds to the exchange of two rows in the determinant, which changes the sign of the determinant.

The advantage of the application of Slater determinants is that matrix elements between them, which are needed to form the CI matrix in Eq. (9.4), can be evaluated in a straightforward way by means of the Slater–Condon rules.

If a Slater determinant is formed from the usual α and β spin-orbitals, then it is also naturally an eigenfunction of the operator $\hat{S}_z$. However, a Slater determinant is an eigenfunction of the total spin operator $\hat{S}^2$ only under the condition that either all electrons are paired or, if unpaired electrons exist, that all of these possess the same spin. Spin-eigenfunctions generally have to be constructed as a linear combination of several Slater determinants.

The construction of spin-eigenfunctions is most readily illustrated in the case of two unpaired electrons placed in two orbitals. There are four possibilities to place two unpaired electrons in two orbitals as illustrated in Figure 9.4. The functions $\Phi_{\alpha\alpha}$ and $\Phi_{\beta\beta}$ are already spin-eigenfunctions as all unpaired electrons possess the same spin. By contrast, linear combinations of $\Phi_{\alpha\beta}$ and $\Phi_{\beta\alpha}$ have

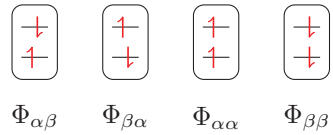


Figure 9.4 Possibilities for placing two unpaired electrons in two orbitals.

to be formed to obtain spin-eigenfunctions. In total the space spanned by the functions shown in Figure 9.4 contains one singlet component

$$|\Phi_S\rangle = 1/\sqrt{2}(|\Phi_{\alpha\beta}\rangle - |\Phi_{\beta\alpha}\rangle) \tag{9.29}$$

and three triplet components with different M_S values

$$|\Phi_{T,+1}\rangle = |\Phi_{\alpha\alpha}\rangle \tag{9.30}$$

$$|\Phi_{T,0}\rangle = 1/\sqrt{2}(|\Phi_{\alpha\beta}\rangle + |\Phi_{\beta\alpha}\rangle) \tag{9.31}$$

$$|\Phi_{T,-1}\rangle = |\Phi_{\beta\beta}\rangle. \tag{9.32}$$

Generally speaking, CI carried out in a Slater determinant basis will lead to correct spin-pure solutions as long as a spin-complete basis constructed from spin-restricted orbitals is used. This follows from the fact that $\hat{H}$ and $\hat{S}^2$ commute and, more specifically, that configurations of different spin do not interact in the CI procedure. Spin contamination only comes into play if unrestricted orbitals are used or if the set of basis functions is not spin-complete, e.g., in the case of spin-flip methods [21]. The practical problem for performing CI in a basis of Slater determinants is not that one obtains incorrect results but that unnecessary basis functions are carried along. For example, in the case of Figure 9.4, one would have to consider, both, $\Psi_{\alpha\beta}$ and $\Psi_{\beta\alpha}$ as basis functions even though it is *a priori* known that only their linear combination Ψ_S will contribute to a singlet wave function. Thus, in a case of two open-shell electrons one would already have to include twice as many configurations as necessary and this factor increases strongly for more open shells.

To overcome the above problem, it is possible to combine several Slater determinants into eigenfunctions of the $\hat{S}^2$ operator. The resulting functions are usually called configuration state functions (CSF). Every CSF is formed as a linear combination of Slater determinants, all possessing the same spatial occupation but different spin. Using CSFs allows the efficiency of CI to be increased by reducing the size of the CI vector. However, it also complicates the CI algorithm. While matrix elements between Slater determinants can simply be evaluated according to the Slater–Condon rules, things become more complicated for CSFs. In fact, much of the algorithmic development in MRCI is concerned with constructing CSFs and computing the matrix elements between them. For this purpose, different approaches such as the symmetric group [20] and the unitary group approach [19] have been developed.

9.2.5 Workflow

A typical workflow for practical uc-MRCI computations is shown in Figure 9.5. The steps involved in the computation of the MRCI energy are shown on the left side, going from bottom to top, whereas the steps for the gradient are on the right side, going from top to bottom. Initially, one has to specify the molecular geometry and basis set, and this information is encoded in the AO integrals. The AO integrals are usually stored on disk. The next crucial step is the generation of the molecular orbitals. This step, discussed in more detail in Section 9.3.4, is most commonly done by MCSCF. To carry out the MCSCF step, the AO integrals are needed as input, and the main parameters for MCSCF are concerned with the active space definition. Subsequently, the MO coefficients

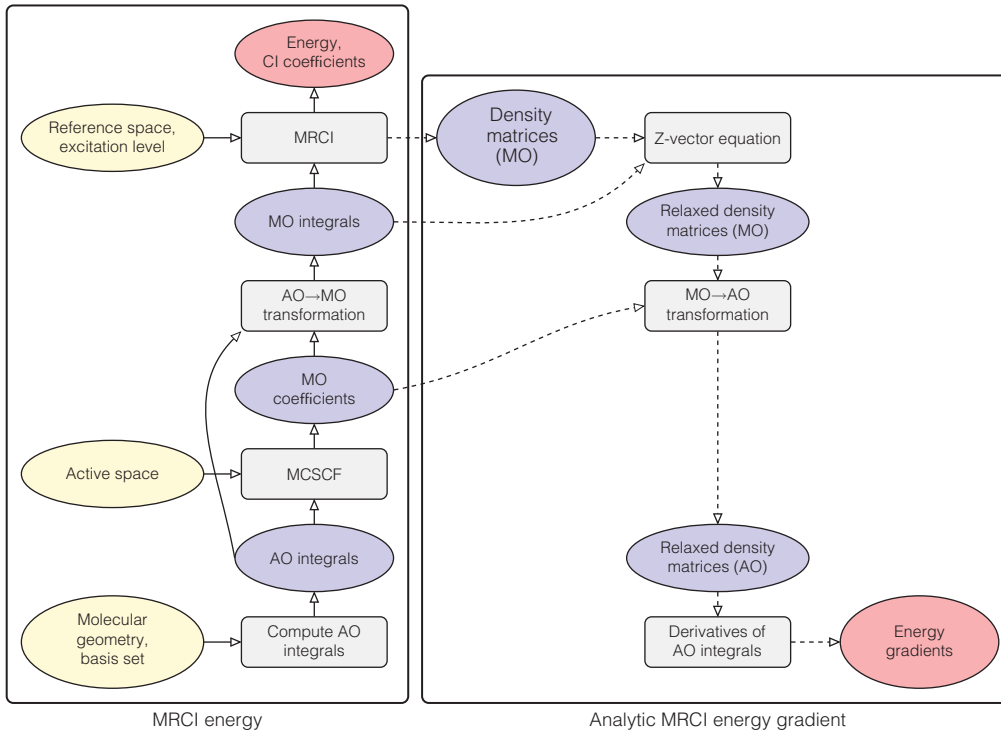


Figure 9.5 Workflow of a typical MRCI computation: Steps for computing the MRCI energy and gradient are shown on the left and right panels, respectively.

are used to transform the AO integrals into the MO basis. The resulting MO integrals are the only type of input data that is required by the MRCI step. The main parameters for the MRCI step are the definition of the reference space and the excitation level. At this point, the matrix eigenvalue equation (9.7) is solved using the MO integrals to construct the $\mathbf{H}$ matrix on-the-fly and the energies E^{α} and CI vectors $\mathbf{c}^{\alpha}$ are obtained as output.

On the right side of Figure 9.5, we also show the computation of energy gradients, which are required for geometry optimizations or on-the-fly dynamics simulations (see chapters 13–17). Due to the variational nature of MRCI and the simple form of its uncontracted variant, it is relatively straightforward to compute analytic gradients based on uc-MRCI and the computational cost is usually only a fraction of the energy calculation. For a detailed discussion on analytic MRCI gradients and the related non-adiabatic coupling vectors we refer readers to Refs [5, 10, 11, 22]. The gradient computation follows similar routes to the energy computation only that it goes the opposite way (starting in the MO basis and finishing in the AO basis) and that the use of AO/MO integrals are exchanged with that of density matrices. In a first step, the density matrices are computed in the MO basis. Subsequently, it is necessary to compute relaxed density matrices that incorporate non-Hellmann–Feynman terms [23], which derive from the fact that the MO-coefficients are not variationally optimized for the MRCI computation, see e.g., Ref. [5]. This step is often denoted the Z-vector equation. The resulting relaxed density matrices are now transformed from the MO to the AO basis. In a final step, the density matrices are contracted with the derivatives of the AO integrals to obtain the energy gradients.

For practical applications, in particular for on-the-fly dynamics simulations, it is important to realize that the workflow proceeds through a number of steps of different scaling behavior and other computational characteristics. For example, the integral computation and transformation steps only depend on the choice of the underlying one-electron basis while the other steps also depend on the many-electron expansion. The MCSCF, MRCI, and Z-vector steps are iterative in nature meaning that the required computational effort can vary between geometries, and the MCSCF and Z-vector steps are both known for causing convergence problems. Depending on the choice of the one-electron basis, active space and excitation level, any one of the steps involved can become the bottleneck of the computation.

9.3 Types of MRCI

In this section, we discuss different variants of MRCI that are used in practical computations. We first discuss the possibility for contracting the CI expansion (Section 9.3.1) to reduce the size of the configuration space. In Section 9.3.2, we continue by specifying different practical methods that allow us to overcome the size-extensivity problem explained in Section 9.2.2. We finish by mentioning two more specific issues: selection schemes for MRCI configuration spaces and the construction of the orbitals underlying the expansion.

9.3.1 Uncontracted and Contracted MRCI

The CI expansion given in Eq (9.1) provides a straightforward and flexible definition for constructing MR wave functions and also leads to a formally simple variational prescription to obtain energy and expansion coefficients by means of Eq. (9.7). This expansion is usually called an uncontracted (uc) MRCI expansion. The main problem of the uc-MRCI expansion is that it can lead to exceedingly high memory requirements and computational cost as the reference space is increased.

The length of the CI vector for uc-MRCI is approximately given as [1]

$$N_{\mathrm{CI}} \approx N_{\mathrm{ref}} \times \binom{n_{\mathrm{el}}}{l} \times \binom{n_{\mathrm{ext}}}{l} \approx \frac{N_{\mathrm{ref}} \times {n_{\mathrm{el}}}^{l} \times {n_{\mathrm{ext}}}^{l}}{(l!)^2} \tag{9.33}$$

where N_{ref} is the number of reference configurations, n_{el} and n_{ext} are the numbers of electrons and external orbitals, and l is the excitation level. This is because for constructing a configuration one has to pick l out of n_{el} electrons and has to place them into l external orbitals, and this has to be done for all N_{ref} reference configurations. In the Davidson algorithm one usually uses a fixed number of subspace expansion vectors, meaning that the memory requirements are proportional to N_{CI}. For practical use cases the computational cost for MRCI scales according to

$$\mathcal{O}(N_{\mathrm{ref}} \times n_{\mathrm{docc}}^{l} \times n_{\mathrm{ext}}^{(l+2)}) \tag{9.34}$$

as explained in Ref. [1]. The problem for, both, memory and computational cost is that these do not just scale with the number of orbitals, as in usual SR computations, but that N_{ref} comes into play as an additional factor. As discussed in Chapter 6, even a moderate reference space such as a CAS(8,8) already contains about 1000 configurations. This means that the memory requirements and the computational cost are increased by a factor of 1000 for uc-MRCI based on this reference space as compared to the SR case. Hence, uc-MRCI can lead to exceedingly high memory requirements and computational cost. This problem can be mitigated by running the calculation in parallel and storing the CI vector in distributed memory [24, 25] allowing computations with over a billion configurations [18], but in many cases it is desirable to reduce the computational effort.

For this reason, contraction schemes have been developed which strongly mitigate the growth of the CI expansion and the concomitant computational cost. One of the most popular contraction schemes is the internal contraction introduced independently by Meyer [26], Werner [27], and Siegbahn [28]. In this approach, taking the electronic ground state as example, the reference configurations are contracted to a fixed function Φ_0 as

$$|\Phi_0\rangle = \sum_{m=1}^{N_{\text{ref}}} c_m^{\text{ref}} |\Phi_m^{\text{ref}}\rangle \tag{9.35}$$

usually substituting the CSF expansion coefficients c_m of a preceding CASSCF/MCSCF calculation for the c_m^{ref} coefficients of Eq (9.35). Single and double excitations are then constructed from this contracted reference wave function by the application of spin-free excitation operators in analogy to the single-reference case only that the reference determinant is replaced by $|\Phi_0\rangle$. In the case of double excitations from the internal orbitals i,j to active or virtual orbitals p,q this results in the functions

$$|\Phi_{ij}^{pq}\rangle = \hat{E}_{ij}^{pq} |\Phi_0\rangle \tag{9.36}$$

where $\hat{E}_{ij}^{pq}$ is a spin-free double-excitation operator. In this equation the internal contraction coefficients are kept fixed. The total internally contracted wave function Ψ is then written as a linear combination

$$|\Psi\rangle = c_0 |\Phi_0\rangle + \sum_{ij}^{\text{int}} \sum_{pq}^{\text{act+virt}} c_{ij}^{pq} |\Phi_{ij}^{pq}\rangle \tag{9.37}$$

where the indices i,j run over internal orbitals and p,q over active plus virtual orbitals. The main advantage achieved by this expansion is the fact that the number of configurations is independent of the number of reference configurations, i.e., one can get rid of the prefactor N_{ref} in Eq. (9.33). The number of variational parameters scales like that of a SRCI expansion, a fact which is responsible for the substantial savings in comparison to the uncontracted MRCI approach.

It should be noted, however, that the internal contraction approach leads to a significantly more involved formalism for the calculation of the matrix elements in comparison to those appearing in Eq. (9.7), since higher-order density matrices are required. Moreover, the internally contracted many-electron basis is non-orthogonal.

This *ic*-MRCISD approach has been successfully developed by Werner and Knowles [29, 30]. In this specific approach, only the double excitations are internally contracted, a procedure which is usually denoted WK or partially contracted (PC) scheme in order to avoid the calculation of reduced density matrices of orders higher than four.

Whereas the calculation of excited states is straightforward in the case of uncontracted MRCI via Eq. (9.7), the situation is more complicated in the case of internally contracted MRCI. In contrast to the uncontracted case, the reference wave function is state-specific. Generalizing Eq. (9.35) and using for the contraction coefficients the MCSCF expansion coefficients for each reference state α separately yields:

$$|\Phi_0^{(\alpha)}\rangle = \sum_{m=1}^{N_{\text{ref}}} c_m^{\text{ref},(\alpha)} |\Phi_m^{\text{ref}}\rangle \tag{9.38}$$

Thus, the double excitations of Eq. (9.36) inherit this state-dependence by defining

$$|\Phi_{ij,(\alpha)}^{pq}\rangle = \hat{E}_{ij}^{pq} |\Phi_0^{(\alpha)}\rangle \tag{9.39}$$

where $\hat{E}_{ij}^{pq}$ is a spin-summed double excitation operator.

A straightforward but computationally expensive method, denoted as the multi-state approach [31], is to take the union of all internally contracted configurations formed from all separate reference functions for all states α up to and including a predefined state β of interest. The wave function for state α is then written as a linear combination of the reference configurations Φ_I, the single excitations Φ_S^a and the double excitations $\Phi_{ij,\gamma}^{ab}$

$$\Psi^{(\alpha)} = \sum_I c_I^{(\alpha)} \Phi_I + \sum_S \sum_a [c_S^{(\alpha)}]_a \Phi_S^a + \sum_{\gamma=1}^{\beta} \sum_{i \geq j} \sum_{ab} [c_{ij,\gamma}^{(\alpha)}]_{ab} \Phi_{ij,\gamma}^{ab}. \tag{9.40}$$

The index I runs over internal configurations, S and a denote the internal and external indices of the single excitations; i, j and a, b denote internal and external orbital indices, respectively. Within this configuration space a variational calculation is performed as outlined in [31].

A computationally simpler method has also been described [31], in which only one state is calculated at a time. In this direct CI procedure, the excited-state eigenvector $\Psi^{(\alpha)}$ is expanded in terms of subspace vectors $\Delta\Psi_\lambda^{(\alpha)}$ as

$$\Psi^{(\alpha)} = \sum_I c_I^{(\alpha)} \Phi_I + \sum_\lambda d_\lambda^{(\alpha)} \Delta\Psi_\lambda^{(\alpha)} \tag{9.41}$$

The index λ counts the subspace iterations. The subspace vectors $\Delta\Psi_\lambda^{(\alpha)}$ are written as

$$\Delta\Psi_\lambda^{(\alpha)} = \sum_S \sum_a [\Delta c_S^{(\alpha,\lambda)}]_a \Phi_S^a + \sum_{i \geq j} \sum_{ab} [\Delta C_{ij}^{(\alpha,\lambda)}]_{ab} \Phi_{ij}^{ab} \tag{9.42}$$

and the expansion coefficients are computed from perturbation theory. Note that only contracted functions for state α are involved. For the βth state the lowest β eigensolutions $\Psi^{(\alpha)}, \alpha = 1, \ldots, \beta$ are calculated and the root most similar to the reference state is selected for $\Psi^{(\alpha)}$ and then the subspace iterations are performed.

One problem encountered in this procedure is root-flipping since the expansion functions for a desired higher excited root might describe a lower-lying root less efficiently and, thus, could be located energetically lower. To circumvent this problem, a projection operator method has been developed in which a modified Hamiltonian matrix is constructed where the lower eigensolutions are removed. For the βth state, the following projection matrix is defined:

$$\mathbf{P}^{(\beta)} = \mathbf{1} - \sum_{\alpha=1}^{\beta-1} \mathbf{c}^{(\alpha)} \mathbf{c}^{(\alpha)\dagger} \tag{9.43}$$

and used for the construction of the corresponding projected Hamiltonian matrix

$$\mathbf{H}^{(\beta)} = \mathbf{P}^{(\beta)} \mathbf{H} \mathbf{P}^{(\beta)} \tag{9.44}$$

$\mathbf{c}^{(\alpha)}$ represents the eigenvector of the αth state, which is the lowest eigenvector of $\mathbf{H}^{(\alpha)}$. Because of the contracted CI, construction of the full projector is not feasible. Instead, an approximate vector is constructed where only such configurations are chosen for $\mathbf{c}^{(\alpha)}$ of Eq. (9.43) which are included in the configuration set of both the αth and βth state calculation.

Finally, all wave functions $\Psi^{(\beta)}, \beta \leq \alpha$, are used to solve the non-orthogonal eigenvalue problem of $H_{\beta\alpha} = \langle \Psi^{(\beta)} | \hat{H} | \Psi^{(\alpha)} \rangle, S_{\beta\alpha} = \langle \Psi^{(\beta)} | \Psi^{(\alpha)} \rangle$. It was found that this approach gave almost identical results to the above-mentioned multi-state method.

9.3.2 MRCI with Extensivity Corrections

As outlined in Section 9.2.2, the lack of size-extensivity is a major problem in the practical application of MRCI. For this reason a number of different extensivity corrections have been

introduced. One differentiates between two types of extensivity corrections, *a posteriori* and intrinsic corrections.

The main idea of *a posteriori* extensivity corrections is to add an energetic correction term at the end of an MRCI computation. In the original Davidson correction [32], this term is given as

$$\Delta E_{\mathrm{DC}} = (1 - c_0^2)\Delta E \tag{9.45}$$

where ΔE is the correlation energy and c_0^2 is the combined weight of the reference configurations

$$c_0^2 = \sum_{k=1}^{N_{\mathrm{ref}}} (c_k^{\mathrm{ref}})^2 = \sum_{k=1}^{N_{\mathrm{ref}}} \langle \Psi | \Phi_k^{\mathrm{ref}} \rangle^2. \tag{9.46}$$

Subsequently, a number of related *a posteriori* corrections have been suggested, mostly based on ΔE, c_0^2, and the number of electrons, see Ref. [3] for additional information. The advantage of *a posteriori* corrections is that these are computationally cheap and numerically stable. The main disadvantage of these methods is that they are no longer variational. This has two important consequences: (i) the extensivity corrected energy is no longer an upper bound to the true energy and (ii) it is not readily possible to compute the gradient of the extensivity corrected energy or to compute extensivity corrections to other wave function properties.

As a second option, intrinsic extensivity corrections have been suggested, which modify the iterative procedure itself. Two prominent procedures in this context are the averaged coupled pair functional (ACPF) [33] and the averaged quadratic coupled cluster (AQCC) methods [34]. In both methods the correlation energy is written as a functional of the CI vector $\mathbf{c}$ using the following form

$$F(\mathbf{c}) = \frac{\left\langle \sum_{m=1}^{N_{\mathrm{CI}}} c_m \Phi_m | \hat{H} - E_{\mathrm{ref}} | \sum_{m=1}^{N_{\mathrm{CI}}} c_m \Phi_m \right\rangle}{\sum_k (c_k^{\mathrm{int}})^2 + G \sum_l (c_l^{\mathrm{ext}})^2} \tag{9.47}$$

where E_{ref} is the energy of the reference wave function. The summations in the denominator go over the internal and external configurations, cf. Eq (9.27). The G value determines the method where the following values are commonly used:

- MRCI: $G = 1$
- MR-ACPF: $G = 2/n_{\mathrm{el}}$
- MR-AQCC: $G = 1 - \dfrac{(n_{\mathrm{el}} - 3)(n_{\mathrm{el}} - 2)}{n_{\mathrm{el}}(n_{\mathrm{el}} - 1)}$

Note that in the case of $G = 1$, the functional simply turns into the MRCI expectation value of Eq. (9.10) whereas non-trivial results are obtained for MR-ACPF and MR-AQCC. It can be shown that the vector minimizing $F(\mathbf{c})$ is an eigenvector of a shifted CI matrix where the energies of all external configurations are lowered by a value of $(1 - G)\Delta E$. Practically, MR-ACPF and MR-AQCC implementations proceed by solving the eigenvalue problem (9.7) with the difference that the energies of all external configurations are shifted using the current estimate of the correlation energy. Upon convergence, this procedure leads to the MR-ACPF/MR-AQCC energy. The CI coefficients are optimized variationally during the MR-ACPF and MR-AQCC procedures and, as a consequence, a generalized Hellmann–Feynman theorem applies, which means that gradients and other properties can be readily computed [35]. However, the obtained energies are not upper bounds to the true energies, i.e., Eqs (9.13)–(9.15) do not hold. In fact convergence of the equations is not guaranteed at all [25], which could lead to a situation where the energy tends toward negative infinity. This occurs because, as the Hamiltonian matrix elements are shifted down, the correlation energy increases, which in turn shifts the Hamiltonian matrix elements further. This problem is similar to the intruder-state problem of CASPT2, see Chapter 10. A second problem deriving from the more complicated form is that the computation of excited states is not as straightforward anymore.

Whereas in the original MRCI method excited states are computed as eigenvectors of the same matrix as the ground state, things become more complicated for MR-ACPF and MR-AQCC as the different states have different correlation-energy dependent shifts. This complicates the computation of transition moments. To overcome this problem, the linear-response theory variant of AQCC (LRT-AQCC) has been developed [36]. This operates by inserting the ground-state correlation energy into Eq. (9.47) independently of the state.

In summary, both types of corrections and different variants of them are in common use. One cannot make a clear preference and the choice of correction will depend on the molecular system under study and on the properties of interest.

9.3.3 Types of Selection Schemes

Selection schemes provide important tools to adjust the often extremely large size of the MRCI expansion. General restrictions such as partitioning the orbital spaces according to RASSCF/GASSCF schemes have already been mentioned in Chapter 6. The occupation restriction of the virtual orbital space is another efficient measure already discussed above. Mostly, excitations of only one or two electrons will be allowed from the internal into the virtual orbital space. The first-order interacting space is a convenient restriction to reduce the amount of double excitations [37, 38]. It includes all configurations which have a non-zero Hamiltonian matrix element with one of the members of the reference space. Single excitations are usually excluded from this restriction.

In many cases the selection of the number and types of active orbitals will be performed based on "chemical intuition" or on experience with previous, similar calculations. Systematic extensions of such orbital spaces are in most cases restricted to smaller molecules. Interesting investigations have been made by Pulay and Hamilton [39] based on the natural orbitals of an unrestricted Hartree–Fock (UHF) calculation to determine recommended sizes of active orbitals, which has led later on to the UNO-CAS procedure [40].

A different approach for reducing the number of configurations is to analyze the energetic contribution of individual configurations and to discard them based on a given energy threshold. Such approaches have been applied successfully already in the early days of CI calculations by Buenker and Peyerimhoff [41, 42] in their MRD-CI program [43]. Extrapolation methods are used to estimate full CI limits. Other advanced computational algorithms have been developed by Hanrath and Engels in their DIESEL-MR-CI program [44]. A similar philosophy embedded in an iterative approach has been used in the difference-dedicated CI procedure of Malrieu and co-workers [45]. The idea has been further refined by Neese leading to the spectroscopically oriented CI method [46].

9.3.4 Construction of Orbitals

The availability of adequate orbitals is a prerequisite for the MRCI calculation and is crucial to the result. There are no formal requirements with respect to the orbitals used in MRCI and a number of different choices are in common use. Note that this is different to methods based on many-body perturbation theory where stricter conditions apply usually, e.g., that the orbitals diagonalize the Fock matrix.

In standard applications, MCSCF calculations are performed for the generation of orbitals. They can be performed at the CASSCF or RASSCF levels, depending on the requirements. When excited state calculations are performed, one usually (but not necessarily always) chooses to apply state-averaging in order to obtain a combined set of MOs that facilitates the computation of transition moments and non-adiabatic couplings. While state-averaging is a common and useful procedure, it is worth pointing out that – at least in a variational sense – the description of each

individual state becomes worse the more states are included in the procedure. This problem can be partially alleviated through performing MRCI on top of the state-averaged orbitals. The MRCI protocol not only introduces correlations through double excitations but also allows for a partial relaxation of the orbitals for each specific state through the single excitations.

The use of MCSCF orbitals is not always the best way to proceed. MCSCF calculations could be too costly because of too large active spaces and/or could show convergence problems or multiple-minima solutions. Furthermore, the orbitals optimal for the uncorrelated MCSCF computation are often no longer adequate once dynamic correlation is included. The use of natural orbitals (NOs), generated by a computationally simpler method, has been suggested, for example. One advantage would be that orbital truncation schemes could be applied in a straightforward way based on NO occupation numbers, especially for the virtual orbitals. Frozen natural orbitals [47] have been found to be well-suited. They are constructed by diagonalizing the virtual orbital block of the density matrix separately. Related topics can be found in the IVO-CASCI method [48] and the high-multiplicity NOs [49].

An interesting path for reducing the computation costs for MRCI has been shown by the group of Carter [50, 51] introducing a local correlation (LC) approach based on localized orbitals. This method follows the work of Sæbo and Pulay on the weak pairs (WP) approximation [52]. It restricts the excitations to only those cases where orbitals are in spatial proximity and, therefore, is especially of relevance for molecules of increasingly larger size.

9.4 Popular Implementations

Before concluding, we give a list of some MRCI implementations focusing on codes that are in wide use and can be readily obtained online (either freely or commercially). For this purpose, Table 9.1

Table 9.1 Popular implementations of variants of the MRCI method.

Program package	Type of MRCI	Availability
COLUMBUS	Ab initio, uc-MRCI www.univie.ac.at/columbus	Free
DALTON	General MR computations http://daltonprogram.org	Free
DFT/MRCI	Semi-empirical www.theochem.hhu.de/software/dftci.html	Free
GAMESS	Ab initio, uc-MRCI www.msg.chem.iastate.edu	Free
MOLCAS	Ab initio, uc-MRCI gitlab.com/Molcas	Free
MOLPRO	Ab initio, ic-MRCI www.molpro.net	Commercial
MNDO	Semi-empirical (OM2/MRCI) *Available on request*	Commercial
ORCA	Ab initio, selected MRCI cec.mpg.de/orcadownload	Free
TIGERCI	Ab initio, localized orbitals carter.princeton.edu/research/software	Free

gives an overview of electronic structure codes that feature the MRCI method, shows the main characteristics and provides a link for download.

9.5 Conclusions

The basic principles of the MRCI method and the underlying schemes to construct the space of the many-electron basis used for the expansion have been described. The use of the MR approach requires a certain understanding of the role of the molecular orbitals taking part in the electronic excitation processes. A division into frozen, reference doubly occupied and reference active orbitals facilitates an efficient construction of the reference configurations. A critical analysis of the choice for the different orbital groups is especially important for the uncontracted MRCI because of the drastic increase of computational resources with increasing active spaces, basis sets and molecular size. In this respect, internal contractions have been proven to enhance the computational efficiency drastically. Overall, MRCI is a useful tool, which can be applied for a large variety of problems, but which probably is most attractive and useful in case of truly difficult problems such as computing energy surfaces for excited states and interactions of different electronic states.

References

1 Sherrill, C.D. and Schaefer, H.F. (1999). The configuration interaction method: advances in highly correlated approaches. *Adv. Quantum Chem.* 34: 143–269.
2 Helgaker, T., Jorgensen, P., and Olsen, J. (2000). *Molecular Electronic-Structure Theory*. Wiley.
3 Szalay, P.G., Müller, T., Gidofalvi, G. et al. (2012). Multiconfiguration self-consistent field and multireference configuration interaction methods and applications. *Chem. Rev.* 112: 108–181.
4 Roos, B.O., Lindh, R., Malmqvist, P. et al. (2016). *Multiconfigurational Quantum Chemistry*. Wiley.
5 Lischka, H., Nachtigallová, D., Aquino, A.J.A. et al. (2018). Multireference approaches for excited states of molecules. *Chem. Rev.* 118: 7293–7361.
6 Roos, B.O. and Siegbahn, P.E.M. (1977). The direct configuration interaction method from molecular integrals. In: *Methods of Electronic Structure Theory*, Modern Theoretical Chemistry (ed. H.F. Schaefer III), 277–318. New York: Springer.
7 Hylleraas, E.A. and Undheim, B. (1930). Numerische berechnung der 2S-terme von ortho- und par-helium. *Zeitschrift für Phys.* 65: 759–772.
8 Jensen, F. (2007). *Introduction to Computational Chemistry*, 2e. Wiley.
9 Page, M., Saxe, P., Adams, G.F., and Lengsfield, B.H. (1984). Multireference CI gradients and MCSCF second derivatives. *J. Chem. Phys.* 81: 434–439.
10 Shepard, R. (1987). Geometrical energy derivative evaluation with MRCI wave functions. *Int. J. Quantum Chem.* 31: 33–44.
11 Lischka, H., Dallos, M., Szalay, P.G. et al. (2004). Analytic evaluation of nonadiabatic coupling terms at the MR-CI level. I. Formalism. *J. Chem. Phys.* 120: 7322–7329.
12 Bartlett, R.J. (1981). Many-body perturbation theory and coupled cluster theory for electron correlation in molecules. *Annu. Rev. Phys. Chem.* 32: 359–401.
13 Shavitt, I. and Bartlett, R.J. (2009). *Many – Body Methods in Chemistry and Physics*. Cambridge: Cambridge University Press.
14 Lyakh, D.I., Musiał, M., Lotrich, V.F., and Bartlett, R.J. (2012). Multireference nature of chemistry: the coupled-cluster view. *Chem. Rev.* 112: 182–243.

15 Pople, J.A., Binkley, J.S., and Seeger, R. (2009). Theoretical models incorporating electron correlation. *Int. J. Quantum Chem.* 10: 1–19.

16 Szabo, A. and Ostlund, N.S. (1996). *Modern Quantum Chemistry*, 1e. Dover Publications, Inc.

17 Dreuw, A. and Head-Gordon, M. (2005). Single-reference ab initio methods for the calculation of excited states of large molecules. *Chem. Rev.* 105: 4009–4037.

18 Plasser, F., Pašalić, H., Gerzabek, M.H. et al. (2013). The multiradical character of one- and two-dimensional graphene nanoribbons. *Angew. Chem. Int. Ed.* 52: 2581–2584.

19 (a) Paldus, J. (1974). Group theoretical approach to the configuration interaction and perturbation theory calculations for atomic and molecular systems. *J. Chem. Phys.* 61: 5321–5330.
(b) Shavitt, I. (1981). The graphical unitary group approach and its application to direct configuration interaction calculations. In: *The Unitary Group for the Evaluation of Electronic Energy Matrix Elements*, Lecture Notes in Chemistry, vol. 22 (ed. J. Hinze), 51–99. Berlin: Springer.

20 Duch, W. and Karwowski, J. (1985). Symmetric group approach to configuration interaction methods. *Comput. Phys. Rep.* 2: 93–170.

21 Krylov, A.I. (2008). Equation-of-motion coupled-cluster methods for open-shell and electronically excited species: the hitchhikers guide to fock space. *Annu. Rev. Phys. Chem.* 59: 433–462.

22 Shepard, R. (1995). The Analytic gradient method for configuration interaction wave functions. In: *Modern Electronic Structure Theory* (ed. D. Yarkony), 345. Singapore: World Scientific.

23 Handy, N.C. and Schaefer, H.F. (1984). On the evaluation of analytic energy derivatives for correlated wave functions. *J. Chem. Phys.* 81: 5031–5033.

24 Dachsel, H., Lischka, H., Shepard, R. et al. (1997). A massively parallel multireference configuration interaction program: the parallel COLUMBUS program. *J. Comput. Chem.* 18: 430–448.

25 Müller, T. (2009). Large-scale parallel uncontracted multireference-averaged quadratic coupled cluster: the ground state of the chromium dimer revisited. *J. Phys. Chem. A* 113: 12729–12740.

26 Meyer, W. (1977). Configuration expansion by means of pseudonatural orbitals. In: *Methods Electronic Structure Theory* (ed. H.F. Schaefer), 413–446. Springer.

27 Werner, H. and Reinsch, E. (1982). The selfconsistent electron pairs method for multiconfiguration reference state functions. *J. Chem. Phys.* 76: 3144–3156.

28 Siegbahn, P.E.M. (1980). Direct configuration interaction with a reference state composed of many reference configurations. *Int. J. Quantum Chem.* 18: 1229–1242.

29 Werner, H.J. and Knowles, P.J. (1988). An efficient internally contracted multiconfiguration-reference configuration interaction method. *J. Chem. Phys.* 89: 5803–5814.

30 Knowles, P.J. and Werner, H.J. (1988). An efficient method for the evaluation of coupling coefficients in configuration interaction calculations. *Chem. Phys. Lett.* 145: 514–522.

31 Knowles, P.J. and Werner, H.J. (1992). Internally contracted multiconfiguration-reference configuration interaction calculations for excited states. *Theor. Chim. Acta* 84: 95–103.

32 Langhoff, S.R. and Davidson, E.R. (1974). Configuration interaction calculations on the nitrogen molecule. *Int. J. Quantum Chem.* 8: 61–72.

33 Gdanitz, R.J. and Ahlrichs, R. (1988). The averaged coupled-pair functional (ACPF): a size-extensive modification of MR CI(SD). *Chem. Phys. Lett.* 143: 413–420.

34 Szalay, P.G. and Bartlett, R.J. (1993). Multi-reference averaged quadratic coupled-cluster method: a size-extensive modification of multi-reference CI. *Chem. Phys. Lett.* 214: 481–488.

35 Lischka, H., Shepard, R., Pitzer, R.M. et al. (2001). High-level multireference methods in the quantum-chemistry program system COLUMBUS: analytic MR-CISD and MR-AQCC gradients and MR-AQCC-LRT for excited states, GUGA spin-orbit CI and parallel CI density. *Phys. Chem. Chem. Phys.* 3: 664–673.

36 Szalay, P.G., Müller, T., and Lischka, H. (2000). Excitation energies and transition moments by the multireference averaged quadratic coupled cluster (MR-AQCC) method. *Phys. Chem. Chem. Phys.* 2: 2067–2073.

37 Bunge, A. (1970). Electronic wave functions for atoms. III. Partition of degenerate spaces and ground state of C. *J. Chem. Phys.* 53: 20–28.

38 Bender, C.F. and Schaefer, H.F. (1971). Electronic splitting between the $^2B_1$ and $^2A_1$ states of the NH_2 radical. *J. Chem. Phys.* 55: 4798–4803.

39 Pulay, P. and Hamilton, T.P. (1988). UHF natural orbitals for defining and starting MC-SCF calculations. *J. Chem. Phys.* 88: 4926–4933.

40 Bofill, J.M. and Pulay, P. (1989). The unrestricted natural orbital-complete active space (UNO-CAS) method: an inexpensive alternative to the complete active space-self-consistent-field (CAS-SCF) method. *J. Chem. Phys.* 90: 3637–3646.

41 Buenker, R.J. and Peyerimhoff, S.D. (1974). Individualized configuration selection in CI calculations with subsequent energy extrapolation. *Theor. Chim. Acta* 35: 33–58.

42 Buenker, R.J., Peyerimhoff, S.D., and Butscher, W. (1978). Applicability of the multi-reference double-excitation CI (MRD-CI) method to the calculation of electronic wave functions and comparison with related techniques. *Mol. Phys.* 35: 771–791.

43 Krebs, S. and Buenker, R.J. (1995). A new tabledirect configuration interaction method for the evaluation of Hamiltonian matrix elements in a basis of linear combinations of spin-adapted functions. *J. Chem. Phys.* 103: 5613–5629.

44 Hanrath, M. and Engels, B. (1997). New algorithms for an individually selecting MR-CI program. *Chem. Phys.* 225: 197–202.

45 García, V.M., Castell, O., Caballol, R., and Malrieu, J.P. (1995). An iterative difference-dedicated configuration interaction. Proposal and test studies. *Chem. Phys. Lett.* 238: 222–229.

46 Neese, F. (2003). A spectroscopy oriented configuration interaction procedure. *J. Chem. Phys.* 119: 9428–9443.

47 Barr, T.L. and Davidson, E.R. (1970). Nature of the configuration-interaction method in ab initio calculations. I. Ne ground state. *Phys. Rev. A* 1: 644–658.

48 Potts, D.M., Taylor, C.M., Chaudhuri, R.K., and Freed, K.F. (2001). The improved virtual orbital-complete active space configuration interaction method, a "packageable" efficient ab initio many-body method for describing electronically excited states. *J. Chem. Phys.* 114: 2592–2600.

49 Lu, Z. and Matsika, S. (2013). High-multiplicity natural orbitals in multireference configuration interaction for excited state potential energy surfaces. *J. Phys. Chem. A* 117: 7421–7430.

50 Walter, D. and Carter, E.A. (2001). Multi-reference weak pairs local configuration interaction: efficient calculations of bond breaking. *Chem. Phys. Lett.* 346: 177–185.

51 Walter, D., Venkatnathan, A., and Carter, E.A. (2003). Local correlation in the virtual space in multireference singles and doubles configuration interaction. *J. Chem. Phys.* 118: 8127–8139.

52 Sæbo, S. and Pulay, P. (1993). Local treatment of electron correlation. *Annu. Rev. Phys. Chem.* 44: 213–236.

10

Multi-Configurational Reference Perturbation Theory with a CASSCF Reference Function

Roland Lindh and Ignacio Fdez. Galván

Department of Chemistry – BMC, Uppsala University, SE-751 23 Uppsala, Sweden

Abstract

The purpose of the present chapter is to give students a detailed introduction to time-independent multi-configurational reference perturbation theory (MRPT), in particular the CASPT2 method, one of the standard tools for the study of excited states of molecular systems of small to intermediate size. To achieve this, in as closed a form as possible, we believe that the chapter should begin with the very basics of Rayleigh–Schrödinger (RS) and Møller–Plesset (MP) perturbation theory (PT) before we present the MRPT. The multi-configurational version of perturbation theory has many features and flaws which are either a direct consequence of the original Rayleigh–Schrödinger formulation, or due to the choice of the zeroth order Hamiltonian. These flaws or peculiarities have to be clearly identified and understood before we proceed toward the more elaborate multi-configurational reference perturbation theories. This is then followed by a section dedicated to the most popular versions of single-state multi-configurational reference perturbation theory. Here we again present the formulation of these methods and also address the various problems they encounter. Multi-state versions of perturbation theory are today based on effective Hamiltonian approximations. Hence, these deserve their own treatment due to the fact that this approach introduces its own kind of features and problems, which need to be addressed separately. At the very end of this chapter we summarize and present an outlook on the subject of MRPT.

Let us briefly, before we start, address the reason behind the need of a perturbation theory based on "multi-configurational reference" functions. Single-configuration reference methods, such as standard coupled-cluster methods and Møller–Plesset perturbation theory, have been successful in modeling molecular systems close to their equilibrium structure. However, at molecular geometries at which such reference functions are qualitatively wrong, as in the case of bond breaking, for example, single-configuration reference perturbation or coupled-cluster theory has shown convergence problems. These approaches frequently produce significant artifacts at such instances. This effect has been demonstrated to be associated with the qualitative incorrectness of the reference functions, from which coupled-cluster or perturbation theory will not be able to recover. Hence, the need for a theory based on reference functions which are qualitatively correct – multi-configurational reference functions. It is not uncommon to find the term "multi-reference" perturbation theory used to mean the same kind of theory. However, it should not be confused with quasi-degenerate perturbation theory, that simultaneously treats

Quantum Chemistry and Dynamics of Excited States: Methods and Applications,
First Edition. Edited by Leticia González and Roland Lindh.

several functions (which can be single- or multi-configurational) as their reference, and we avoid this usage, except when it is part of a specific method's name.

10.1 Rayleigh–Schrödinger Perturbation Theory

This section will cover a brief introduction to perturbation theory as presented originally by Schrödinger [1], who based it on previous work by Lord Rayleigh [2]. We will develop the theory explicitly up to second order, extensions to higher order are trivial. We will introduce the standard equation for the perturbed wave functions and the corresponding energies, but we will will also introduce the so-called $2n + 1$ rule for the efficient computation of energies. This section will be split into two sections: one for the basics of the single-state theory and a section on the issue of convergence properties and intruder states. The former is consistent with the conventional presentation of the subject, while the latter is often not discussed in the depth that students would appreciate. The second section deals with an important problem which arises in both single- and multi-configurational reference PTs. The discussion of the foundations will be resumed with quasi-degenerate perturbation theory after the most significant single-state multi-configurational reference perturbation methods have been presented.

10.1.1 The Single-State Theory

We would first like to point out that the formulation introduced by Schrödinger does not make any explicit reference to the model used for describing wave functions – single configuration, multiple configurations, or something else – but is basically a thought experiment based on the notion that exact solutions (energies and wave functions) are known for a particular Hamiltonian

$$\hat{H}_0 \Psi_i^{(0)} = E_i^{(0)} \Psi_i^{(0)}, \tag{10.1}$$

these energies and wave functions are denoted as unperturbed reference energies and wave functions of state i, the "(0)" is a notation that we will use to indicate the unperturbed entities. We also note that, due to the Hamiltonian being a Hermitian operator, the corresponding eigenfunctions form a set of complete functions, i.e., any other function can be expressed as a linear combination of these eigenfunctions. We will then try to find out to what extent this information can be used to solve the Schrödinger equation for any other Hamiltonian – a perturbed Hamiltonian – which can be expressed as

$$\hat{H} = \hat{H}_0 + \lambda V, \tag{10.2}$$

provided that V, the perturbation potential, is "small". In the context of this chapter, and of RS theory, V is a time-independent perturbation. λ is here introduced as the perturbation parameter, and when it takes the values 0 and 1 this Hamiltonian is equal to the unperturbed Hamiltonian and to the fully perturbed or "target" Hamiltonian, respectively. At this point we note that the theory is independent of what the actual nature of the partitioning and the unperturbed reference functions is, so these are completely arbitrary, apart from the desire that the resulting perturbation expansion be rapidly converging such that the approach will be useful in practice. There will now follow two different ways to derive the theory – the conventional and the variational approach.

10.1.1.1 The Conventional Projectional Derivation

The perturbation parameter is instrumental in our pursuit to make any sense of how the information of the unperturbed system can be of any use. In passing we note that RS perturbation theory is based on a *global* partitioning of the Hamiltonian, something it does not share with

subsequent implementations of perturbation theory. The theory then proceeds with introducing a Taylor expansions of the energy and the wave functions in terms of the perturbation parameter and with the unperturbed energies and wave functions as the reference points of the expansions. That is,

$$E_i = E_i^{(0)} + E_i^{(1)}\lambda + E_i^{(2)}\lambda^2 + \dots \tag{10.3}$$

for the energies and

$$\Psi_i = \Psi_i^{(0)} + \Psi_i^{(1)}\lambda + \Psi_i^{(2)}\lambda^2 + \dots \tag{10.4}$$

for the wave function, where $E_i^{(k)}$ and $\Psi_i^{(k)}$ correspond to the kth-order correction to the perturbed energy and wave function, respectively, as a function of the perturbation. At this time we do not need to discuss what these corrections are, apart the obvious fact that the $E_i^{(k)}$ are the kth-order derivatives of the energy with respect to the perturbation parameter λ. In the following we will make a difference between the nth-order corrected energy and the nth-order correction of the energy. While the latter is $E^{(n)}$, as found in Eq. (10.3), the former is the sum of the unperturbed energy and all corrections up to the nth order:

$$E_i^{[n]} = \sum_{i=0}^{n} E_i^{(i)}. \tag{10.5}$$

The RS theory now proceeds with expressing the Schrödinger equation of the perturbed system in terms of the unperturbed Hamiltonian, energies and wave functions, that is,

$$\begin{aligned}(\hat{H}_0 + \lambda V)(\Psi_i^{(0)} + \Psi_i^{(1)}\lambda + \Psi_i^{(2)}\lambda^2 + \dots) \\ = (E_i^{(0)} + E_i^{(1)}\lambda + E_i^{(2)}\lambda^2 + \dots)(\Psi_i^{(0)} + \Psi_i^{(1)}\lambda + \Psi_i^{(2)}\lambda^2 + \dots)\,.\end{aligned} \tag{10.6}$$

After reordering the terms in the equation we find that

$$\begin{aligned}&(\hat{H}_0\Psi_i^{(0)} - E_i^{(0)}\Psi_i^{(0)})\lambda^0 \\ &+ (\hat{H}_0\Psi_i^{(1)} + V\Psi_i^{(0)} - E_i^{(0)}\Psi_i^{(1)} - E_i^{(1)}\Psi_i^{(0)})\lambda^1 \\ &+ (\hat{H}_0\Psi_i^{(2)} + V\Psi_i^{(1)} - E_i^{(0)}\Psi_i^{(2)} - E_i^{(1)}\Psi_i^{(1)} - E_i^{(2)}\Psi_i^{(0)})\lambda^2 \\ &+ \dots = 0.\end{aligned} \tag{10.7}$$

The Schrödinger equation is fulfilled for any value of the perturbation parameter, λ. Hence, this implies that each row of the equation above is independently identical to zero, and we have the following equations explicitly up to second order of the perturbation,

$$\hat{H}_0\Psi_i^{(0)} = E_i^{(0)}\Psi_i^{(0)} \tag{10.8}$$

$$\hat{H}_0\Psi_i^{(1)} + V\Psi_i^{(0)} = E_i^{(0)}\Psi_i^{(1)} + E_i^{(1)}\Psi_i^{(0)} \tag{10.9}$$

$$\hat{H}_0\Psi_i^{(2)} + V\Psi_i^{(1)} = E_i^{(0)}\Psi_i^{(2)} + E_i^{(1)}\Psi_i^{(1)} + E_i^{(2)}\Psi_i^{(0)}, \tag{10.10}$$

and then in general,

$$(\hat{H}_0 - E_i^{(0)})\Psi_i^{(n)} = -V\Psi_i^{(n-1)} + \sum_{k=1}^{n} E_i^{(k)}\Psi_i^{(n-k)}. \tag{10.11}$$

We identify the first equation as the Schrödinger equation of the unperturbed system. The subsequent equations, however, leave us far from our goal – expressions of the corrections to the perturbed energies and wave functions defined completely in terms of entities of the unperturbed system. We note that Eq. (10.11) is invariant to any admixture of the reference state, $\Psi_i^{(0)}$, into the

nth-order correction, $\Psi_i^{(n)}$, since this will leave the the left-hand side (LHS) unchanged. We relieve this ambiguity by the introduction of the *intermediate normalization*:

$$\langle \Psi_i^{(0)} | \Psi_i^{(k)} \rangle = \delta_{0,k} \ . \tag{10.12}$$

The combination of Eqs. (10.11) and (10.12) leads to two important results. First, Eq. (10.11) completely defines the nth-order correction of the wave function from lower-order corrections. Here we note a recursive nature of Eq. (10.11), a non-recursive variant would lead to an expression in terms of the Bloch equation – more about this later. Second, the exact energy can be computed as

$$E_i = \langle \Psi_i^{(0)} | \hat{H} | \Psi_i \rangle \ . \tag{10.13}$$

Equations (10.9) to (10.11), and so forth, can now be projected from the LHS and integrated with the unperturbed wave functions. We start by projecting Eq. (10.11) with $\Psi_i^{(0)}$ this gives

$$\langle \Psi_i^{(0)} | (\hat{H}_0 - E_i^{(0)}) | \Psi_i^{(n)} \rangle = -\langle \Psi_i^{(0)} | V | \Psi_i^{(n-1)} \rangle + \sum_{k=1}^{n} \langle \Psi_i^{(0)} | E_i^{(k)} | \Psi_i^{(n-k)} \rangle \ . \tag{10.14}$$

Using the Hermiticity $\hat{H}_0$ and the intermediate normalization this expression reduces to

$$E_i^{(n)} = \langle \Psi_i^{(0)} | V | \Psi_i^{(n-1)} \rangle \ . \tag{10.15}$$

This seems to mean that we need the $(n-1)$th-order correction to the wave function in order to compute the nth-order correction to the energy. This is, however, far from true and will be discussed below.

Let us now return to the corrections of the wave function and Eq. (10.11). We will rearrange the equation to form our general vehicle for the computation of corrections to the wave function. We first note that the equation is devoid of any component of $\Psi_i^{(0)}$. This property was actually used when going from Eq. 10.14 to Eq. 10.15, since

$$\begin{aligned} \langle \Psi_i^{(0)} | (\hat{H}_0 - E_i^{(0)}) | \Psi_i^{(n)} \rangle &= \langle \Psi_i^{(0)} | \hat{H}_0 | \Psi_i^{(n)} \rangle - \langle \Psi_i^{(0)} | E_i^{(0)} | \Psi_i^{(n)} \rangle \\ &= E_i^{(0)} \delta_{0,n} - E_i^{(0)} \delta_{0,n} = 0. \end{aligned} \tag{10.16}$$

This means that we can now multiply Eq. (10.11) with the inverse of $(\hat{H}_0 - E_i^{(0)})$ avoiding any singularity problems. We get

$$\Psi_i^{(n)} = (\hat{H}_0 - E_i^{(0)})^{-1} \left(-V\Psi_i^{(n-1)} + \sum_{k=1}^{n} E_i^{(k)} \Psi_i^{(n-k)} \right) \ . \tag{10.17}$$

Knowing that the correction and any term of Eq. (10.15) are devoid of $\Psi_i^{(0)}$, we introduce the projection operator

$$\hat{Q} = 1 - |\Psi_i^{(0)}\rangle\langle\Psi_i^{(0)}| \ . \tag{10.18}$$

We can now simplify Eq. (10.17) without any approximation and get

$$\Psi_i^{(n)} = \hat{Q}(\hat{H}_0 - E_i^{(0)})^{-1} \hat{Q} \left(-V\Psi_i^{(n-1)} + \sum_{k=1}^{n-1} E_i^{(k)} \Psi_i^{(n-k)} \right) \ . \tag{10.19}$$

Note that we have reduced the summation range, the projection operator will now, whenever needed, cancel any appearance of the reference wave function. Let us pause at this point and see what we have accomplished. We have first derived our master equation (10.11). From this equation we have derived a general expression for the correction to the energy, see Eq. (10.15), in which the nth-order correction to the energy requires the $(n-1)$th-order correction to the wave function.

Finally, in Eq. (10.19) we now have a recursive approach in which corrections to the wave function can be computed. We also note that in RSPT the corrected energy is neither an eigenvalue nor an expectation value, and the corrected wave function is not an eigenfunction to the Hamiltonian.

Before we complete this section let us return to our dangling question: "To which order do we need to know the correction of the wave function if we want to compute the nth-order correction to the energy?" From what we have seen so far the answer seems to be $(n-1)$. However, as noted before, this is wrong. Consider that the first- and second-order corrections to the wave function, according to Eq. (10.19), can be expressed as

$$\Psi_i^{(1)} = -\hat{Q}(\hat{H}_0 - E_i^{(0)})^{-1}\hat{Q}V\Psi_i^{(0)}, \tag{10.20}$$

and

$$\Psi_i^{(2)} = \hat{Q}(\hat{H}_0 - E_i^{(0)})^{-1}\hat{Q}(-V + E_i^{(1)})\Psi_i^{(1)}. \tag{10.21}$$

We can then, for example, write the third-order correction of the energy as

$$\begin{aligned} E_i^{(3)} &= \langle\Psi_i^{(0)}|V|\Psi_i^{(2)}\rangle \\ &= \langle\Psi_i^{(0)}|V\hat{Q}(\hat{H}_0 - E_i^{(0)})^{-1}\hat{Q}(-V + E_i^{(1)})|\Psi_i^{(1)}\rangle \\ &= \langle\Psi_i^{(1)}|V - E_i^{(1)}|\Psi_i^{(1)}\rangle \ . \end{aligned} \tag{10.22}$$

This clearly demonstrates that the third-order energy can be computed from the first-order corrections to the wave function, a first demonstration of Wigner's $2n+1$ rule. To develop this in more depth is beyond the scope of this chapter. For those students interested in this we recommend the excellent book by Helgaker and co-workers [3], where this particular issue is analyzed with a variational Lagrangian approach.

Up to this point we have not been specific about the corrections to the wave function, other than the fact that we use the intermediate normalization. At this point we will make an ansatz based on our observation that the corrections are orthogonal to the reference functions, that is, we will take the set of unperturbed eigenfunctions as a basis set,

$$\Psi_i^{(1)} = \sum_{j\neq i} c_{ji}\Psi_j^{(0)}. \tag{10.23}$$

Using our master equation, Eq. (10.11), replacing the nth-order correction with the previous ansatz, left-multiplying with $\langle\Psi_l^{(0)}|$ and finally integrating we get

$$\sum_{j\neq i} c_{ji}\langle\Psi_l^{(0)}|(\hat{H}_0 - E_i^{(0)})|\Psi_j^{(0)}\rangle = -\langle\Psi_l^{(0)}|V|\Psi_i^{(0)}\rangle \ , \tag{10.24}$$

which simplifies to

$$c_{li} = \frac{\langle\Psi_l^{(0)}|V|\Psi_i^{(0)}\rangle}{E_i^{(0)} - E_l^{(0)}} \ . \tag{10.25}$$

Following Eq. (10.15) and substituting the first-order correction as described above, we can formulate the second-order correction to the energy as

$$E_i^{(2)} = \sum_{j\neq i} \frac{\langle\Psi_i^{(0)}|V|\Psi_j^{(0)}\rangle\langle\Psi_j^{(0)}|V|\Psi_i^{(0)}\rangle}{E_i^{(0)} - E_j^{(0)}} \ . \tag{10.26}$$

At this point we stop any further analysis. There are several reasons for this. Two dominating ones are that the expressions get more and more complicated and cumbersome, but also, and maybe more importantly, perturbation theory beyond second-order corrections to the energy has not been

found to constitute a significant enough improvement to compensate the additional complexity and computational expense. The curious readers can either derive the next set of equations themselves or find them published elsewhere.

10.1.1.2 The Bi-Variational Approach

In the above conventional approach we projected the Schrödinger equation from the left with the reference wave function, $\Psi_i^{(0)}$. This introduces an asymmetry which is in particular manifested by the shape of the equation in which we compute corrections to the energy (see Eq. 10.15). In this section we will explore an alternative derivation which will be in accordance with Wigner's $2n+1$ rule. To explore this we start from an expression which is originally symmetric with respect to its bra and ket side. For convenience, in the rest of this section we will omit the subindex i in the wave functions and energies, since all expressions refer to a single state. In order to maintain the symmetry we will use a bi-functional expression of a generalization of the expectation value for the energy, that is

$$\mathcal{E}(\lambda, \bar{\Psi}, \tilde{\Psi}) = \langle \bar{\Psi} | \hat{H}(\lambda) | \tilde{\Psi} \rangle \,, \tag{10.27}$$

which we note is equivalent to the variational expression of the exact energy,

$$\mathcal{E}(\lambda, \Psi, \Psi) = E(\lambda, \Psi),$$

given an arbitrary actual value of the perturbational parameter, and if the normalized wave functions $\bar{\Psi}$ and $\tilde{\Psi}$ are equivalent to the exact eigenfunction, Ψ, of the Hamiltonian $\hat{H}(\lambda)$, in which case we can also write

$$\langle \Psi | \hat{H}(\lambda) - \mathcal{E}(\lambda, \Psi, \Psi) | \Psi \rangle = 0. \tag{10.28}$$

We note that, $\mathcal{E}(\lambda, \Psi, \Psi)$, the exact energy, is at the global minimum and the gradient, with respect to any variation of the wave function, is zero:

$$\begin{aligned} \nabla_{\bar{\Psi}} \mathcal{E}|_{\bar{\Psi}=\Psi} &= 0 \\ \nabla_{\tilde{\Psi}} \mathcal{E}|_{\bar{\Psi}=\Psi} &= 0. \end{aligned} \tag{10.29}$$

Since this holds for any value of the perturbational parameter, we also have that the derivatives must vanish order by order, for $k > 0$:

$$\begin{aligned} \nabla_{\bar{\Psi}} \mathcal{E}^{(n)}|_{\bar{\Psi}=\Psi} &= 0 \qquad & \nabla_{\bar{\Psi}^{(k)}} \mathcal{E}^{(n)}|_{\bar{\Psi}=\Psi} &= 0 \\ \nabla_{\tilde{\Psi}} \mathcal{E}^{(n)}|_{\bar{\Psi}=\Psi} &= 0 \qquad & \nabla_{\tilde{\Psi}^{(k)}} \mathcal{E}^{(n)}|_{\bar{\Psi}=\Psi} &= 0. \end{aligned} \tag{10.30}$$

These conditions of vanishing gradients for the variational parameters will help us later in simplifying the equations we derive. From now on, we will consider the case where $\bar{\Psi} = \tilde{\Psi} = \Psi$, but we will continue to make a difference between the two functions, $\bar{\Psi}$ and $\tilde{\Psi}$, although we set them to be identical, Ψ.

We proceed from Eq. (10.28) and replace Ψ with $\Psi^{(0)} + \hat{Q}\Psi$

$$\begin{aligned} &\langle \Psi^{(0)} | \hat{H} - \mathcal{E} | \Psi^{(0)} \rangle + \langle \Psi^{(0)} | \hat{H} - \mathcal{E} | \hat{Q}\Psi \rangle \\ &+ \langle \hat{Q}\Psi | \hat{H} - \mathcal{E} | \Psi^{(0)} \rangle + \langle \hat{Q}\Psi | \hat{H} - \mathcal{E} | \hat{Q}\Psi \rangle = 0. \end{aligned} \tag{10.31}$$

After expanding $\hat{H} = \hat{H}_0 + \lambda V$ and some rearrangement, noting that $\langle \Psi^{(0)} | Q\Psi \rangle = 0$, we have

$$\begin{aligned} \mathcal{E}(\lambda, \Psi, \Psi) = \langle \Psi^{(0)} | \hat{H}_0 | \Psi^{(0)} \rangle &+ \lambda \langle \Psi^{(0)} | V | \Psi^{(0)} \rangle \\ &+ \lambda \langle \Psi^{(0)} | V | \hat{Q}\Psi(\lambda) \rangle + \lambda \langle \hat{Q}\Psi(\lambda) | V | \Psi^{(0)} \rangle \\ &+ \langle \hat{Q}\Psi(\lambda) | \hat{H}_0 + \lambda V - \mathcal{E}(\lambda, \ldots) | \hat{Q}\Psi(\lambda) \rangle \,, \end{aligned} \tag{10.32}$$

where the third and fourth terms have not been merged together to reflect the difference of the bra and ket wave functions. It could be argued that nothing has been gained by the introduction of this equation, compared with a conventional approach using a projection with the reference wave function on the Schrödinger equation. In particular, the exact energy appears on both sides of the equation and it is more complicated. However, there is a significant important advantage with this approach as will become evident soon. Moreover, an analysis order by order shows that the contribution from the energy on the right-hand side (RHS) will only depend on lower-order corrections to the energy than the LHS and will not involve the actual value of the exact energy. That is, at the stationary point at which $\mathcal{E} = E_{\text{exact}}$, we have that

$$\begin{aligned}\mathcal{E}^{(n)} = E^{(n)} = \delta_{n0}E^{(0)} &+ \sum_{k=0}^{n-1}\langle\Psi^{(k)}|V|\Psi^{(n-1-k)}\rangle \\ &+ \sum_{k=1}^{n-1}\langle\Psi^{(k)}|\hat{H}_0 - E^{(0)}|\Psi^{(n-k)}\rangle \\ &- \sum_{k=1}^{n-2}E^{(k)}\sum_{l=1}^{n-k-1}\langle\Psi^{(l)}|\Psi^{(n-k-l)}\rangle \ . \end{aligned} \tag{10.33}$$

In particular, we note that this formulation is completely symmetric with respect to the treatment of the bra and ket sides. At this point we stop using the explicit reference to $\mathcal{E}$ and will just use E. However, we should keep in mind that when we differentiate we do that on $\mathcal{E}$! To continue, Eq. (10.33) now trivially provides us with expressions for the corrected energy up to first order which are identical to those obtained from the projectional derivation,

$$E^{(0)} = \langle\Psi^{(0)}|\hat{H}_0|\Psi^{(0)}\rangle \tag{10.34}$$

$$E^{(1)} = \langle\Psi^{(0)}|V|\Psi^{(0)}\rangle \ , \tag{10.35}$$

However, at second order things take a different direction, we get

$$E^{(2)} = \langle\Psi^{(0)}|V|\Psi^{(1)}\rangle + \langle\Psi^{(1)}|V|\Psi^{(0)}\rangle + \langle\Psi^{(1)}|\hat{H}_0 - E^{(0)}|\Psi^{(1)}\rangle \tag{10.36}$$

instead of (cf. Eq. 10.15) just $E^{(2)} = \langle\Psi^{(0)}|V|\Psi^{(1)}\rangle$! For the third-order correction we do not get $E^{(3)} = \langle\Psi^{(0)}|V|\Psi^{(2)}\rangle$, but

$$\begin{aligned}E^{(3)} = &\langle\Psi^{(0)}|V|\Psi^{(2)}\rangle + \langle\Psi^{(1)}|V|\Psi^{(1)}\rangle \\ &+ \langle\Psi^{(2)}|V|\Psi^{(0)}\rangle + \langle\Psi^{(1)}|\hat{H}_0 - E^{(0)}|\Psi^{(2)}\rangle \\ &+ \langle\Psi^{(2)}|\hat{H}_0 - E^{(0)}|\Psi^{(1)}\rangle - E^{(1)}\langle\Psi^{(1)}|\Psi^{(1)}\rangle \ . \end{aligned} \tag{10.37}$$

What is going on? Well, remember that these expressions are variational. Taking the second-order energy correction, for example, differentiating with respect to $\bar{\Psi}^{(1)}$, and equating to zero we find that

$$\nabla_{\bar{\Psi}^{(1)}}E^{(2)} = V|\Psi^{(0)}\rangle + (\hat{H}_0 - E^{(0)})|\Psi^{(1)}\rangle = 0, \tag{10.38}$$

and similarly for $\tilde{\Psi}^{(1)}$. That allows us to reduce Eq. (10.36) in two different ways getting,

$$E^{(2)} = \langle\Psi^{(0)}|V|\Psi^{(1)}\rangle = \langle\Psi^{(1)}|V|\Psi^{(0)}\rangle \ . \tag{10.39}$$

To continue, for the third-order correction to the energy, we find that the gradient reveals the following conditions,

$$\nabla_{\bar{\Psi}^{(1)}}E^{(3)} = (\hat{H}_0 - E_0)|\Psi^{(2)}\rangle + (V - E^{(1)})|\Psi^{(1)}\rangle = 0 \tag{10.40}$$

$$\nabla_{\bar{\Psi}^{(2)}}E^{(3)} = V|\Psi^{(0)}\rangle + (\hat{H}_0 - E^{(0)})|\Psi^{(1)}\rangle = 0, \tag{10.41}$$

and the equivalent equations for the bra side. Here we are left with some ambiguity: Which of the relations should we use to eliminate redundancies and in which order? We will, of course, use the conditions that eliminate the higher-order corrections first, and continue until we have exhausted all possibilities. For the third-order energy correction, for example, we note that if we use one of the last conditions (e.g., for the bra) first and then the corresponding one of the first conditions, we will gain nothing – we will simply regenerate the Eq. (10.15) expression, albeit in a very clumsy way. However, using the two last conditions we compile the third-order correction to the energy as

$$E^{(3)} = \langle \Psi^{(1)} | V - E^{(1)} | \Psi^{(1)} \rangle \,, \tag{10.42}$$

where we have completely eliminated $\Psi^{(2)}$ from the expression!

We will now generalize this procedure to reduce Eq. (10.33), for which we examine all elements of the form

$$\langle \Psi^{(k)} | \hat{O} | \Psi^{(l)} \rangle \,,$$

that is,

$$\langle \Psi^{(k)} | V | \Psi^{(n-1-k)} \rangle$$
$$\langle \Psi^{(k)} | \hat{H}_0 - E^{(0)} | \Psi^{(n-k)} \rangle$$
$$\langle \Psi^{(m)} | \Psi^{(n-k-m)} \rangle \,.$$

To tackle this reduction let us rewrite Eq. (10.33) in a grid form, representing the $\langle \Psi^{(k)} | \hat{O} | \Psi^{(l)} \rangle$ terms. The indices k and l will be used to label columns and rows, and each cell contains the associated operator $\hat{O}$ (empty cells are zero). We will display cells in green if they survive the reduction process.

Starting with $n = 1$, we get

l \ k	0
0	V

where the only cell represents $\langle \Psi^{(0)} | V | \Psi^{(0)} \rangle$. To proceed with $n = 2$, we have

l \ k	0	1
0		V
1	V	$\hat{H}_0 - E^{(0)}$

The gradient conditions are now seen as full rows or columns, and we can eliminate one of them, indicated with a zigzag strike-out. The result corresponds to $\langle \Psi^{(0)} | V | \Psi^{(1)} \rangle$. For $n = 3$ we again see the gradient conditions as rows and columns, and we can strike out $k = 2$ and $l = 2$:

l \ k	0	1	2
0			V
1		$V - E^{(1)}$	$\hat{H}_0 - E^{(0)}$
2	V	$\hat{H}_0 - E^{(0)}$	

It is only allowed to strike out a row or column if it contains no non-empty cell that has already been stricken out. We could have stricken out $k = 2$ and $k = 1$ to generate $\langle\Psi^{(0)}|V|\Psi^{(2)}\rangle$, but instead we have here obtained $\langle\Psi^{(1)}|V - E^{(1)}|\Psi^{(1)}\rangle$, which requires only first-order entities (cf. Eq. 10.22). Similarly, for $n = 4$

l \ k	0	1	2	3
0				V
1		$-E^{(2)}$	$V - E^{(1)}$	$\hat{H}_0 - E^{(0)}$
2		$V - E^{(1)}$	$\hat{H}_0 - E^{(0)}$	
3	V	$\hat{H}_0 - E^{(0)}$		

and for $n = 5$

l \ k	0	1	2	3	4
0					V
1		$-E^{(3)}$	$-E^{(2)}$	$V - E^{(1)}$	$\hat{H}_0 - E^{(0)}$
2		$-E^{(2)}$	$V - E^{(1)}$	$\hat{H}_0 - E^{(0)}$	
3		$V - E^{(1)}$	$\hat{H}_0 - E^{(0)}$		
4	V	$\hat{H}_0 - E^{(0)}$			

As for $n = 3$, we could eliminate all columns $k > 0$ (or rows $l > 0$), and we would end up with Eq. (10.15). But in order to eliminate higher-order entities, we proceed by eliminating the rows and columns with the highest k and l indices until elimination is impossible. By inspection we find that

$$E^{(4)} = \langle\Psi^{(2)}|V - E^{(1)}|\Psi^{(1)}\rangle - E^{(2)}\langle\Psi^{(1)}|\Psi^{(1)}\rangle \tag{10.43}$$

$$E^{(5)} = \langle\Psi^{(2)}|V - E^{(1)}|\Psi^{(2)}\rangle - 2E^{(2)}\langle\Psi^{(1)}|\Psi^{(2)}\rangle - E^{(3)}\langle\Psi^{(1)}|\Psi^{(1)}\rangle\,. \tag{10.44}$$

Thus, a general grid representation gives us the following table.

l \ k	0	1	2	$\cdots$	i	$\cdots$	$n-3$	$n-2$	$n-1$
0									V
1		$-E^{(n-2)}$	$-E^{(n-3)}$		$-E^{(n-i-1)}$		$-E^{(2)}$	$V - E^{(1)}$	$\hat{H}_0 - E^{(0)}$
2		$-E^{(n-3)}$	$-E^{(n-4)}$		$-E^{(n-i-2)}$		$V - E^{(1)}$	$\hat{H}_0 - E^{(0)}$	
$\vdots$									
i		$-E^{(n-i-1)}$	$-E^{(n-i-2)}$		$V - E^{(1)}$				
$\vdots$									
$n-3$		$-E^{(2)}$	$V - E^{(1)}$						
$n-2$		$V - E^{(1)}$	$\hat{H}_0 - E^{(0)}$						
$n-1$	V	$\hat{H}_0 - E^{(0)}$							

First, it is obvious from the elimination process that all rows and columns that contain the operator $\hat{H}_0 - E^{(0)}$ can be ignored. At the end of the elimination process we will be left with a block of elements of the size $m \times (m-1)$ if $n = 2m$, the even case, or $m \times m$ if $n = 2m+1$, the odd case – the green blocks in all table representations above. We note that in all terms but one the operator is of the nature $-E^{(o)}$ and that o has a value such that, together with the order indices of the wave function corrections, adds up to n ($k + l + o = n$). The remaining element is

$$\langle \Psi^{(m)} | V - E^{(1)} | \Psi^{(m+i-1)} \rangle$$

if we generalize the odd and even case to $n = 2m + i$, where i is 0 or 1 for the even and odd cases, respectively. Summing up the remaining elements row-/column-wise we derive the following general expression:

$$n = 2m + i, \quad i \in \{0, 1\}$$
$$E^{(n)} = \langle \Psi^{(m)} | V | \Psi^{(m+i-1)} \rangle - \sum_{k=1}^{m} \sum_{l=1}^{m+i-1} E^{(n-k-l)} \langle \Psi^{(k)} | \Psi^{(l)} \rangle \, . \tag{10.45}$$

This expression can be found in ref. [3] – in a derivation based on a variational analysis of a Lagrange multiplier approach – albeit a different order of summation. To conclude this section, the primary benefit of this formulation was the derivation of a variational expression for the second-order correction to the total energy – a feature some developers use for finding the correction in an efficient way. Additionally, if higher-order corrections are to be computed this formalism offers an optimal procedure for that purpose.

10.1.2 Convergence Properties and Intruder States

The use of perturbation theory seems at first glance to be a rock-solid approach that would never fail. However, this is unfortunately not the case. Here we will discuss the convergence properties of perturbation theory to some extent and the associated issue of intruder states. This short section will not only clarify the issue but will also suggest a remedy to correct the problem. The presentation here, based on Kato's work [4], will be simplified and brief. A longer and more detailed presentation can be found in refs. [5] and [3].

In the text that follows we will address the issue of the convergence properties of RSPT. We will do the core analysis based on a simple two-state case – a study of a multi-state case would make the analysis complicated and messy. We will initially do that on the assumption that the perturbation strength parameter λ takes only real values. This will give us insight into how the relative values of the energy gap between the two states and the coupling element play a fundamental role in determining whether or not the perturbation series of a particular partitioning will converge. We will then introduce a generalization of this by allowing the perturbation strength to take complex values – this will improve and clarify the analysis. In combination with this we will introduce the notion of degeneracies, which can be characterized as the origin of the so-called *front-* or *back-door intruder states*. Finally, we will, in the two-state model, introduce a *gap-shift parameter* which will let us model and modify the conditions of perturbation expansion convergence and the appearance of intruder states. The gap-shift technique will also be the inspiration for techniques to improve the convergence and remove intruder states.

Assume that we, in a two-state model, partition the Hamiltonian as follows:

$$\boldsymbol{H} = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} + \lambda \begin{pmatrix} 0 & \delta \\ \delta & 0 \end{pmatrix} \, , \tag{10.46}$$

where α and β represent the eigenvalues of the zeroth-order states, δ corresponds to the interaction term between the reference states over the perturbation potential, and λ is the standard perturbational strength parameter. For this trivial case we can compute the exact energies of the final states as

$$E_{\pm} = \frac{\alpha+\beta}{2} \pm \frac{\sqrt{(\beta-\alpha)^2+4\lambda^2\delta^2}}{2} \,. \tag{10.47}$$

We rewrite this for the ground state as

$$E = \frac{\alpha+\beta}{2} - \frac{|\beta-\alpha|}{2}\sqrt{1+\frac{4\lambda^2\delta^2}{(\beta-\alpha)^2}}, \tag{10.48}$$

and note that the way forward to reshape this into an RS summation order by order goes via the Taylor expansion of $\sqrt{1+x}$ for which we know that the expansion converges if $|x| < 1$. This leads, as a requirement for convergence, to the inequality

$$\frac{4\delta^2}{(\beta-\alpha)^2} < 1 \tag{10.49}$$

and we get the following final convergence condition:

$$|\beta-\alpha| > 2|\delta|, \tag{10.50}$$

that is, the energy gap between the states should be larger than twice the interaction between the unperturbed states over the perturbation potential to guarantee convergence in a perturbational approach. This is now a tool that will assist us in using information from the unperturbed states. Clearly, it suggests that we have a serious problem if two unperturbed states are close to degenerate.

To proceed, we will also have to analyze what happens when we have a degeneracy for the perturbed states. To do this we will use the technique of Kato [4]. That is, we follow the same procedure as above with the two new ingredients, (*i*) λ is allowed to take complex values, $\lambda = \lambda_{\mathrm{R}} + \mathrm{i}\lambda_{\mathrm{I}}$, and (*ii*) we identify values for λ at which the two states are degenerate, $E_i(\lambda_{\mathrm{R}} + \mathrm{i}\lambda_{\mathrm{I}}) = E_j(\lambda_{\mathrm{R}} + \mathrm{i}\lambda_{\mathrm{I}})$. In passing it is noted that these degeneracies occur in conjugate pairs (λ, λ^*).

In this analysis it is found that the convergence of the perturbation series depends on the behavior of the energies $E_i(\lambda)$ inside the unit circle $|\lambda| \leq 1$ in the complex plane. That is, if the value of λ that causes a degeneracy in the perturbed states represents a point inside the unit circle, we have a convergence issue and the state that becomes degenerate with the reference state is called an intruder state. The intruder states are divided into two classes, (*i*) front-door intruders ($\lambda_{\mathrm{R}} > 0$), which in real life tend to correspond to low-excitation states, and (*ii*) back-door intruders ($\lambda_{\mathrm{R}} < 0$), corresponding to high-excitation states. The difference between the front- and back-door intruders is manifested in different convergence patterns of the energy (oscillating versus alternating), change in reference weight (slow versus abrupt), etc.

To get a better understanding of this situation an alternative partitioning is utilized,

$$\boldsymbol{H} = \begin{pmatrix} \alpha & 0 \\ 0 & \beta+\gamma \end{pmatrix} + \lambda \begin{pmatrix} 0 & \delta \\ \delta & -\gamma \end{pmatrix} \,, \tag{10.51}$$

where $\gamma \in \mathbb{R}$ is the gap-shift parameter. The solutions in this case are given by

$$E_{\pm} = \frac{\alpha+\beta+(1-\lambda)\gamma}{2} \pm \frac{\sqrt{[|\beta-\alpha|+(1-\lambda)\gamma]^2+4\lambda^2\delta^2}}{2} \,. \tag{10.52}$$

We note that this particular partitioning is such that at $\lambda = 1$ the energy does not depend on the gap-shift parameter. Equation (10.52) has a degeneracy (as can be found by setting the expression

in the square root to zero) for a value of the perturbation strength parameter of

$$\lambda_{\pm} = \frac{|\beta - \alpha| + \gamma}{4\delta^2 + \gamma^2}(\gamma \pm \mathrm{i}2\delta), \tag{10.53}$$

and if $|\lambda_{\pm}| < 1$ we have a convergence problem and a corresponding intruder state. We can twist this analysis around and reformulate it to answer the question "What is the minimal gap-shift parameter which will make the perturbation expansion converge?", arriving at the result

$$\gamma_c = \frac{4\delta^2 - (\beta - \alpha)^2}{2|\beta - \alpha|} \,. \tag{10.54}$$

Alternatively, we can paraphrase Eq. (10.50) to see the influence of the gap-shift on the original convergence condition, that is

$$2|\delta| < |\beta - \alpha|\sqrt{1 + \frac{2\gamma}{|\beta - \alpha|}}. \tag{10.55}$$

To conclude the analysis, convergence is not an obvious property of the perturbation expansion. The relative sizes of energy differences and interactions over the perturbation are of significance for the convergence condition. We find that when convergence breaks down this is associated with front- or back-door intruder states. It was finally demonstrated that the convergence issue could be managed by introducing a gap-shift parameter which, with the appropriate value, will guarantee convergence of the perturbation expansion. We will now continue with a presentation of two methods that will solve convergence issues in perturbation theory (as implemented in the CASPT2 method): the real and imaginary shift methods – both draw their inspiration from the gap-shift parameter trick.

10.1.2.1 Real and Imaginary Shift Techniques

In the gap-shift technique suggested in association with multi-state situations the partitioning of the Hamiltonian is established as

$$\hat{H} = (\hat{H}_0 + \epsilon\hat{Q}) + (V - \epsilon\hat{Q}), \tag{10.56}$$

where ϵ is the gap-shift (also referred to by some as a level-shift) parameter – a small real number. With this partitioning we get that the energy corrected up to first order is invariant to the gap-shift parameter,

$$\begin{aligned} E^{[1]} &= E^{(0)} + E^{(1)} \\ &= \langle \Psi^{(0)} | \hat{H}_0 + \epsilon\hat{Q} | \Psi^{(0)} \rangle + \langle \Psi^{(0)} | V - \epsilon\hat{Q} | \Psi^{(0)} \rangle \\ &= \langle \Psi^{(0)} | \hat{H}_0 | \Psi^{(0)} \rangle + \langle \Psi^{(0)} | V | \Psi^{(0)} \rangle \,. \end{aligned} \tag{10.57}$$

However, after this we will start to see effects from the gap-shift parameter, for example, the coefficients of the first-order correction to the wave function under the influence of the gap-shift parameter are computed as (see Eq. 10.25)

$$\tilde{c}_{ji} = -\frac{\langle \Psi_j^{(0)} | V | \Psi_i^{(0)} \rangle}{E_j^{(0)} - E_i^{(0)} + \epsilon} = \frac{\langle \Psi_j^{(0)} | V | \Psi_i^{(0)} \rangle}{E_i^{(0)} - E_j^{(0)} - \epsilon} \,, \tag{10.58}$$

and the second-order energy correction, which also depends on the gap-shift parameter, is expressed as

$$\tilde{E}_i^{(2)} = \sum_{j \neq i} \frac{\langle \Psi_i^{(0)} | V | \Psi_j^{(0)} \rangle \langle \Psi_j^{(0)} | V | \Psi_i^{(0)} \rangle}{E_i^{(0)} - E_j^{(0)} - \epsilon} \,, \tag{10.59}$$

where we indicate with a tilde that these quantities are obtained in the presence of the gap shift.

We will try to remove the influence of the gap-shift parameter as if there was no intruder-state problem. This dependence is removed by the following procedure and approximation. First we note that we can, after some struggle, rewrite the level-shifted second-order correction to the energy as

$$\tilde{E}_i^{(2)} = E_i^{(2)} + \epsilon \sum_{j \neq i} \tilde{c}_{ji}^2 \left(1 - \frac{\epsilon}{E_i^{(0)} - E_j^{(0)}} \right), \tag{10.60}$$

which under the assumption $|E_j^{(0)} - E_i^{(0)}| \gg |\epsilon|$ (equivalent to not having an intruder-state problem) gives us the approximation

$$E_i^{(2)} \approx \tilde{E}_i^{(2)} - \epsilon \left(\frac{1}{\tilde{\omega}_i} - 1 \right) \equiv E_i^{(2),\mathrm{LS}}, \tag{10.61}$$

where $\tilde{\omega}_i = \left(\sum_{j \neq i} \tilde{c}_{ji}^2 + 1 \right)^{-1}$ is the weight of the reference function in the normalized expansion of the wave function including the first-order correction under the influence of the gap-shift parameter. As can be seen from the equation, in the absence of any intruder state $\tilde{E}_i^{(2)}$ is linear with respect to the gap-shift parameter, and $E_i^{(2),\mathrm{LS}}$, the level-shift (LS) corrected second-order energy, is, to first order in ϵ, equal to the result of the original approach. Hence, the technique of gap shifting can be used in general, since it will not significantly affect the result when there is no convergence problem. For a detailed description of this type of a real-shift technique and results of using it can be found in the work by Roos and Andersson [6]. In particular, they have demonstrated that the potential energy surface away from the molecular structures at which singularities arise is unaffected by the gap-shift approach and the use of the level-shifted corrected values for the second-order energy corrections gives results which are much more independent of the size of the gap-shift parameter. See fig. 10.1 for a case in which the method is instrumental in avoiding intruder states.

This work inspired Forsberg and Malmqvist [7] to find alternative ways to get a second-order corrected energy which was even more independent of the size of the gap-shift parameter. Forsberg and Malmqvist introduced the imaginary gap-shift parameter, $\mathrm{i}\epsilon$, to investigate this. It should be mentioned that the idea of an imaginary shift had been explored just a year earlier [8]. Similar to the case of a real-valued gap-shift parameter we will have that the energy up to first order is unaffected by the perturbation, the parameter shows up again in the second-order correction to the energy and elimination will have to be made in a fashion different from what was demonstrated above. We get trivially that the second-order energy correction is

$$\tilde{E}_i^{(2)} = \sum_{j \neq i} \frac{\langle \Psi_i^{(0)} | V | \Psi_j^{(0)} \rangle \langle \Psi_j^{(0)} | V | \Psi_i^{(0)} \rangle (E_i^{(0)} - E_j^{(0)} - \mathrm{i}\epsilon)}{(E_i^{(0)} - E_j^{(0)})^2 + \epsilon^2}, \tag{10.62}$$

where we are interested in the real part and want to explicitly eliminate the dependence of the strength of the gap-shift parameter. This would be trivial if the calculation was performed with complex arithmetic, but that is seldom the case and something which we try to avoid. To our assistance comes the formalism we derived using the bi-functional variational technique to derive the energy equations order by order. In particular, we have that the second-order energy correction (Eq. 10.36), can be expressed as (now with the modified $\hat{H}_0$ and V terms)

$$\begin{aligned} &\mathcal{E}_i^{(2)}(\bar{\Psi}_i = \Psi_i, \tilde{\Psi}_i = \Psi_i) \\ &\quad = \langle \Psi_i^{(0)} | V | \Psi_i^{(1)} \rangle + \langle \Psi_i^{(1)} | V | \Psi_i^{(0)} \rangle + \langle \Psi_i^{(1)} | \hat{H}_0 + \mathrm{i}\epsilon - E_i^{(0)} | \Psi_i^{(1)} \rangle , \end{aligned} \tag{10.63}$$

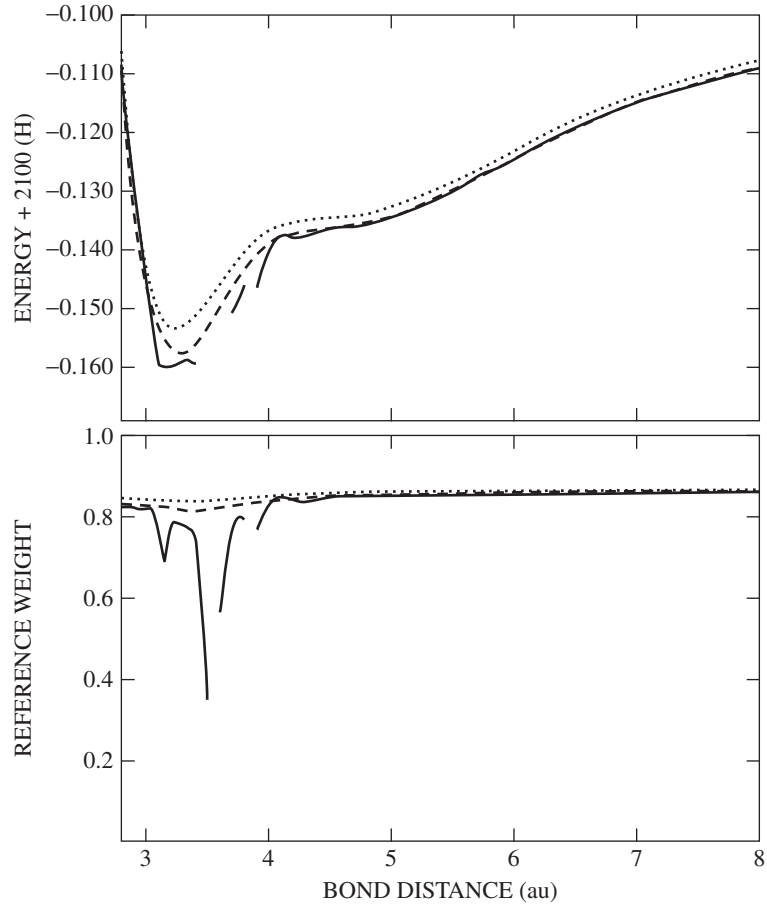


Figure 10.1 The potential energy curve for the ground state of Cr_2 for three values of the level shift, 0.05 E_h (solid line), 0.10 E_h (dashed line), and 0.20 E_h (dotted line). The LS correction of Eq. (10.61) has been applied. The lower diagram gives the corresponding weight of the CASSCF reference function. Reprinted with permission from ref. [6]. Copyright 1995, Elsevier.

where the gap-shift parameter associated with V disappears due to the orthogonality of $\Psi^{(0)}$ and $\Psi^{(1)}$ under intermediate normalization. The derivative of the expression above is

$$\nabla_{\tilde{\Psi}_i^{(1)}} \mathcal{E}_i^{(2)}|_{\tilde{\Psi}_i=\Psi_i,\tilde{\Psi}_i=\Psi_i} = V|\Psi_i^{(0)}\rangle + (\hat{H}_0 + \mathrm{i}\epsilon - E_i^{(0)})|\Psi_i^{(1)}\rangle = 0, \tag{10.64}$$

which we can solve in an iterative manner using a preconditioned conjugate gradient (PCG) procedure (it should be of no surprise that this equation becomes equivalent to Eq. (10.11) if we project with $\hat{Q}$ – we are only interested in the part of the LHS which is perpendicular to the reference function). In the implementation of the PCG a particular routine and the preconditioner are slightly modified to incorporate the imaginary shift. At convergence the trial vector is used in a last call to the routine, but this time for the unshifted $\hat{H}_0$ and the corrected energy is computed using Eq. (10.63). In the closing of their report, Forsberg and Malmqvist list the advantages of the new method compared to the real-shift approach. First, the effect of the imaginary-shift approach on points away from the singularity is smaller than that of the real-shift method (compare figs. 10.1 and 10.2). Second, the imaginary-shift method has the advantage that it does not create new singularities – the real-valued shift technique does – but removes the singularities altogether.

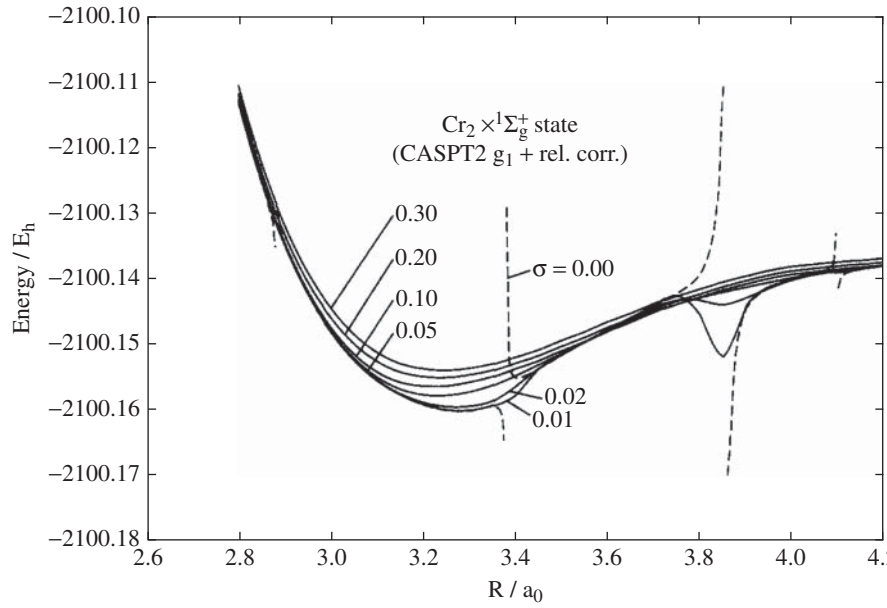


Figure 10.2 The CASPT2 potential energy curves for the ($^1\Sigma_g^+$) ground state of Cr_2, unshifted and with five values of the imaginary shift (called σ here). Reprinted with permission from ref. [7]. Copyright 1997, Elsevier.

This is particularly important for the purpose of studies of excited states. To summarize their conclusions, the imaginary-shift technique includes no calculations using complex arithmetic, the computed energy is real, and the procedure seems to have an edge over real-valued shift techniques in resolving cases with close to perturbation expansion divergence and the associated intruder states. We finally note that the iterative procedure mentioned above is a normal procedure in perturbation theory, although for some cases such as the Møller–Plesset partitioning in association with single determinant wave functions a non-iterative approach is accessible.

We will later see other examples of shift techniques, in particular the sometimes controversial IPEA shift (see section 10.3.3.5). We will have examples in which the shift is not universal but actually specific with respect to the state in the first-order interacting space. This is trivially achieved with a state-specific shift parameter designed as follows

$$\sum_{j\neq i} \epsilon_j |\Psi_j^{(0)}\rangle\langle\Psi_j^{(0)}| \,, \tag{10.65}$$

where the sum runs over the space of the complementary states. We will return to this as we discuss the IPEA shift.

This ends the discussion of single-state perturbation theory from the perspective of RSPT. In the subsequent sections we will see examples of specifics with respect to the partitioning of the Hamiltonian, how to deal with the fact that we do not have a complete reference space, what happens if the partitioning is state dependent as in the case of a multi-state formalism, etc.

10.2 Møller–Plesset Perturbation Theory

Møller–Plesset perturbation theory (MPPT), in which the reference functions are single determinants, is discussed here for several reasons. One, it deals with the issue of what to do if there

is no set of known reference functions which are eigenfunctions to $\hat{H}_0$. Two, CASPT2 theory is designed to coincide with MPPT2 results in the case the complete active space SCF collapses to a single-determinant SCF wave function. Hence, a special study of MPPT has its merits. We will also introduce the use of the second-quantization formalism here, where it is simpler, such that the student will be familiar with the technique once it is used later in the chapter. There exist alternative single-determinant partitioning schemes, such as, for example, the Epstein–Nesbet partitioning [9, 10]. However, the use of these in combination with multi-configurational reference functions is limited and will not be discussed in this chapter.

10.2.1 The Reference Function

The preceding sections leave us in the position of being able to discuss the Møller–Plesset PT without any new twists or approximations as compared to the Rayleigh–Schrödinger approach. However, let us first inspect the implications of our reference function's being a single Slater determinant (SD). This is at this time a bit tedious, but it will help us in understanding the procedure and decisions that will be made at the CASPT2 level of theory. Due to the single-determinant nature of the self-consistent field (SCF) wave function the one- and two-electron particle density matrices, expressed in the basis of the SCF canonical orbitals, reduce to

$$D_{ij} = 2\delta_{ij} \tag{10.66}$$

and

$$\Gamma_{ijkl} = 4\delta_{ij}\delta_{kl} - 2\delta_{il}\delta_{jk} \ , \tag{10.67}$$

and correspondingly the closed-shell SCF energy (an *expectation value*, not an eigenvalue), expressed in spatial doubly occupied orbitals, is compiled as

$$E_{\mathrm{SCF}} = 2\sum_i h_{ii} + \sum_{ij}[2(ii|jj) - (ij|ji)] \ , \tag{10.68}$$

note that we have not added the constant corresponding to the nucleus–nucleus repulsion, V_{NN}, as is conventionally added to the electronic SCF energy. The corresponding orbital Fock operator, is written as

$$\hat{f}_{pq}(\Psi^{(0)}) = \hat{E}_{pq}\left(h_{pq} + \sum_k [2(pq|kk) - (pk|kq)]\right) \tag{10.69}$$

$$= \hat{E}_{pq}(h_{pq} + v_{pq}(\Psi^{(0)})) \ , \tag{10.70}$$

note that the summation over the orbitals k is in terms of the occupied orbitals of the reference SCF wave function $\Psi^{(0)}$ – this part of the equation expresses the interaction of a single electron with all other electrons. Hence, the recursive nature of the operator. This operator will yield the Fock matrix, which on diagonalization produces the canonical orbitals and the eigenvalues ϵ_i – the orbital energy of canonical orbital i. At convergence, the orbitals in the summation, expressing the electron density, are identical to the canonical orbitals – self-consistency is reached.

Let us define a molecular Fock operator as

$$\hat{F}(\Psi^{(0)}) = \sum_{pq} \hat{f}_{pq}(\Psi^{(0)}), \tag{10.71}$$

which has the following expectation value for the restricted closed-shell SCF wave function:

$$\langle \Psi^{(0)}|\hat{F}(\Psi^{(0)})|\Psi^{(0)}\rangle = 2\sum_i \epsilon_i \ , \tag{10.72}$$

where the summation runs over the n doubly occupied spatial orbitals and their associated orbital energy (whereas in Eq. (10.71) the double summation runs over *all* spatial orbitals, note the index usage according to Section 1.5). The orbitals energies, ϵ_i, in the canonical SCF basis is expressed as

$$\epsilon_i = h_{ii} + \sum_j [2(ii|jj) - (ij|ij)]. \tag{10.73}$$

This allows us to write the restricted closed-shell SCF electronic energy as

$$E_{\text{SCF}} = 2\sum_i \epsilon_i - \sum_{ij} [2(ii|jj) - (ij|ji)] \ . \tag{10.74}$$

In the sum of the orbital energies we sum the interaction energies of one electron with all other electrons. This summation will double-count the electron–electron interaction, hence the subtraction of the latter term in the expression above to remove the double counting. Alternatively, we can write the SCF energy without making explicit use of the two-electron integrals:

$$E_{\text{SCF}} = \sum_i (h_{ii} + \epsilon_i). \tag{10.75}$$

To further elaborate on the implications, a given molecular Fock operator – i.e., the interaction field defined by the set of occupied canonical orbitals – has eigenfunctions which are determinants of n orbitals selected out of the full set of occupied and virtual canonical orbitals. We note that there is one determinant which corresponds to the determinant of only the occupied canonical orbitals – the self-consistent solution – while the other determinants, although they certainly are valid eigenfunctions of the given molecular Fock operator, they do not correspond to self-consistent excited states. This is, however, not a requirement or a problem for perturbation theory. We only require that any Hamiltonian operator we use has a set of eigenfunctions which span the full Hilbert space of the system and that we can express them in a systematic and efficient way.

10.2.2 The Partitioning of the Hamiltonian

Now we naturally follow Møller and Plesset, who suggested using as a reference Hamiltonian the molecular Fock operator, $\hat{H}_0 = \hat{F}$, that is, the total Hamiltonian is partitioned as follows,

$$\begin{aligned} \hat{H} &= \hat{h}^{(1)} + \hat{V} \\ &= \hat{F} + \lambda \left(\hat{V} - \sum_{pq} \hat{E}_{pq} v_{pq} \right) \ . \end{aligned} \tag{10.76}$$

By defining $\hat{\mathcal{V}}$ as the term in parentheses and introducing a constant C (to be determined later), we get,

$$\hat{H} = (\hat{F} + C) + \lambda(\hat{\mathcal{V}} - C). \tag{10.77}$$

Note that, while the RHS expression in Eq. (10.76) is general, this partitioning depends on the reference function – the Fock operator is a function of the electron density of the reference state. This trivially gives us the reference energy,

$$E^{(0)}_{\text{MP}} = \langle \Psi^{(0)} | \hat{F} + C | \Psi^{(0)} \rangle = 2\sum_i \epsilon_i + C, \tag{10.78}$$

and the first-order correction of the energy,

$$E^{(1)}_{\text{MP}} = \left\langle \Psi^{(0)} \middle| \hat{V} - \sum_{pq} \hat{E}_{pq} v_{pq} - C \middle| \Psi^{(0)} \right\rangle = \langle \Psi^{(0)} | \hat{\mathcal{V}} | \Psi^{(0)} \rangle - C. \tag{10.79}$$

The energy corrected to first order (recall Eq. 10.5) is identical to the SCF energy,

$$E_{\mathrm{MP}}^{[1]} = E_{\mathrm{MP1}} = \langle \Psi^{(0)} | \hat{F} | \Psi^{(0)} \rangle + C + \langle \Psi^{(0)} | \hat{\mathcal{V}} | \Psi^{(0)} \rangle - C$$
$$= \langle \Psi^{(0)} | \hat{F} + \hat{\mathcal{V}} | \Psi^{(0)} \rangle = E_{\mathrm{SCF}} \, . \tag{10.80}$$

That is, the SCF total energy of the determinant is in general recovered only after the first-order correction and actual corrections to the SCF energy appear first at second order in the perturbation expansion. An alternative approach is to set $C = \langle \Psi^{(0)} | \hat{\mathcal{V}} | \Psi^{(0)} \rangle$, using a shifted-$\hat{H}_0$ operator. In this approach we have that $E^{(0)} = E_{SCF}$ and $E^{(1)} = 0$, and for higher orders of the correction the parameter disappears altogether due to the orthonormality of the reference states – in all, the modification could be considered semantic. In fact, the only significant difference is that the SCF determinants are now eigenfunctions of the shifted-$\hat{H}_0$ operator, something which we will meet in association the description of some of the multi-configurational reference function approaches discussed below.

To continue, the second-order correction to the energy can be computed as

$$E_{\mathrm{MP}}^{(2)} = \langle \Psi^{(0)} | \hat{\mathcal{V}} | \Psi^{(1)} \rangle = \langle \Psi^{(0)} | \hat{H} - \hat{H}_0 | \Psi^{(1)} \rangle \, , \tag{10.81}$$

which suggests that we need to explicitly resolve $\Psi^{(1)}$, something we will do in the following section.

10.2.3 The First-Order Interacting Space and Second-Order Energy Correction

In order to proceed, we need to find the states with which, to first order, the reference function will interact. Before that, however, let us contemplate the impact of this statement. As we explore the effect of the excited states of the unperturbed system on the first-order correction to the wave function, we note that there is a hidden simplification in the equations. In our general ansatz, Eq. (10.23), we imply a summation over *all* excited states. However, upon examination of the equation for the corresponding coefficients of the first-order corrections (see Eq. 10.25) we find that only excited states which interact with the reference state over the perturbation potential have coefficients which are non-zero. We call this space the *first-order interacting space*, a space much smaller than the full Hilbert space. For the SCF reference determinant this translates to the fact that the first-order interacting space contains determinants which at most correspond to double replacements of orbitals from the reference SD.

For a given Fock operator, that is, once we have defined the reference SD which generates the reference electronic density and the operator, we can generate all excited states (eigenfunctions of the unperturbed system) for that particular operator by doing replacements of the occupied orbitals with the virtual (unoccupied) orbitals, starting with one, then two, and so on. As noted above, these "excited states" do not correspond to the true self-consistent excited states of the SCF procedure – that would require us to iteratively redefine the electronic density of the Fock operator to be self-consistent. One would then assume that the first-order correction to $\Psi^{(0)}$ would correspond to such generated eigenfunctions by doing a single excitation – from an occupied to a virtual orbital – that is, to form states as

$$\Psi_i^{a(0)} = \hat{E}_{ai} \Psi^{(0)} \, . \tag{10.82}$$

That would require us to form matrix elements such as

$$\langle \Psi^{(0)} | \hat{H} - \hat{H}_0 | \Psi_i^{a(0)} \rangle = \langle \Psi^{(0)} | \hat{H} | \Psi_i^{a(0)} \rangle = f_{ai} = 0 \tag{10.83}$$

when we determine the coefficients of those states in the first-order correction to the wave function (see Eq. 10.23). However, as seen above, the matrix element is identical to zero: the second term

vanishes since the reference state is an eigenstate to $\hat{H}_0$ while the first term is zero due to the variational minimization of the SCF energy – the Fock matrix is diagonalized such that these elements are identical to zero (Brillouin's theorem).

Let us now (*i*) construct the excited states, (*ii*) obtain their coefficients in the first-order corrected wave function, and (*iii*) subsequently compute the second-order correction of the energy. We do this using both spin-specific and spin-averaged operators, the latter used in the second quantization formalism, as they operate on the reference closed-shell single determinant. In the first case the procedure will be trivial, while in the second case the maths will not be kind to us. The latter will train us for the ordeal that will come when generating excited states for multi-configurational reference functions, a procedure which is always done in the spin-averaged formalism. Here we will have to introduce a new strategy to work around the problem.

We start with the spin-specific case where the replacement operator is expressed as $\hat{a}^\dagger_{a_\sigma}\hat{a}_{i_\sigma}$. In it we will have an operator that annihilates an electron with spin σ in occupied orbital i and places an electron with the same spin in virtual orbital a. For the rest of this subsection, on the use of spin-specific replacement operator, we will use spin-orbitals and index accordingly – that is, no explicit index for the spin. As discussed above the ground-state reference function will not interact with states that correspond to a single replacement – we need double replacements. To proceed, we will represent the doubly excited states as

$$\Psi^{ab}_{ij} = \hat{a}^\dagger_b \hat{a}_j \hat{a}^\dagger_a \hat{a}_i \Psi^{(0)} \,. \tag{10.84}$$

In a setting of using normalized molecular spin-orbitals the replacement operators will generate a new normalized determinant. Moreover, we note that these generated states are orthogonal – a desirable property. Finally, we have to take the four-fold permutational symmetry, $\Psi^{ab}_{ij} = \Psi^{ba}_{ji} = \Psi^{ab}_{ji} = \Psi^{ba}_{ij}$ into account. That is, the first-order wave function is expressed as

$$\Psi^{(1)} = \sum_{\substack{a>b \\ i>j}} c_{ijab} \Psi^{ab}_{ij}, \tag{10.85}$$

where the restricted summation over the spin-orbitals is introduced to include only the unique excited states. The individual coefficients are computed as

$$c_{ijab} = \frac{\langle \Psi^{(0)} | \hat{V} | \Psi^{ab}_{ij} \rangle}{E^{(0)} - E^{ab}_{ij}} \,. \tag{10.86}$$

The denominator trivially reduces to $\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b$ since $E^{ab}_{ij} = E^{(0)} - \epsilon_i - \epsilon_j + \epsilon_a + \epsilon_b$. The numerator is computed to be

$$\langle \Psi^{(0)} | \hat{V} | \Psi^{ab}_{ij} \rangle = (ai|bj) - (aj|bi). \tag{10.87}$$

Thus, the first-order correction to the wave function is expressed as

$$\Psi^{(1)} = \sum_{\substack{a>b \\ i>j}} \frac{(ai|bj) - (aj|bi)}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b} \Psi^{ab}_{ij}. \tag{10.88}$$

We are now in a position to compute the second-order energy correction in terms of spin-orbitals as

$$E^{(2)} = \langle \Psi^{(0)} | \hat{V} | \Psi^{(1)} \rangle = \sum_{\substack{a>b \\ i>j}} \frac{[(ai|bj) - (aj|bi)]^2}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b} \,. \tag{10.89}$$

If we unrestrict the summation, expand the square expression and finally sum over the spin we obtain the second-order correction of the energy in spatial orbitals as

$$\begin{aligned} E^{(2)} &= \frac{1}{4}\sum_{abij}\frac{[(ai|bj)-(aj|bi)]^2}{\epsilon_i+\epsilon_j-\epsilon_a-\epsilon_b} \\ &= \frac{1}{2}\sum_{abij}\frac{(ai|bj)^2}{\epsilon_i+\epsilon_j-\epsilon_a-\epsilon_b} - \frac{1}{2}\sum_{abij}\frac{(ai|bj)(aj|bi)}{\epsilon_i+\epsilon_j-\epsilon_a-\epsilon_b} \end{aligned} \tag{10.90}$$

where we use the fact that the $\sum(ai|bj)^2$ and $\sum(aj|bi)^2$ terms are equal, since they are independent of the order of summation, now we switch to spatial orbital indexation

$$\begin{aligned} E^{(2)} &= 2\sum_{abij}\frac{(ai|bj)^2}{\epsilon_i+\epsilon_j-\epsilon_a-\epsilon_b} - \sum_{abij}\frac{(ai|bj)(aj|bi)}{\epsilon_i+\epsilon_j-\epsilon_a-\epsilon_b} \\ &= \sum_{abij}\frac{(ai|bj)[2(ai|bj)-(aj|bi)]}{\epsilon_i+\epsilon_j-\epsilon_a-\epsilon_b} \ . \end{aligned} \tag{10.91}$$

We have obtained the appropriate equations for the computation of the second-order correction to the energy expressed in both spin-orbitals (see Eq. 10.90) and spatial orbitals (see Eq. 10.91). We did so by using a technique expressed in spin-orbitals, which provided us with (*i*) an easy way to generate orthonormal excited states and (*ii*) trivial equations to compute the corresponding coefficients for the first-order correction to the wave function. In closing this passage, we note that the computation of the second-order correction of the energy in the closed-shell case of MP can be solved in a single step. We do not need to explicitly store the coefficients of the first-order corrected wave function – these are evaluated from analytical integrals which are computed on the fly as we need them. This is a luxury we will not find in the multi-configurational reference function version of the theory – here the coefficients have to be computed in an iterative procedure and need to be explicitly stored before we compile the second-order correction to the energy. It should also be pointed out that the summations in the above equations run over *all* orbitals in the system. It is common in practical calculations to reduce the ranges by ignoring the lowest-lying occupied orbitals and/or the highest-lying virtual orbitals. In the first case – what is usually known as *frozen core* approximation – the inner occupied orbitals are not considered when generating the first-order interacting space, but their Coulomb potential must still be taken into account.

We will now see how we would proceed if we had to repeat these steps using spin-averaged replacement operators in association with a second-quantized Hamiltonian. The procedure will, to some extent, introduce a formalism that is not obviously simpler, however, it will be a general recipe that will guarantee that the generated states have the same eigenvalues of the spin operators $\hat{S}^2$ and $\hat{S}_z$ as the reference functions, since $\hat{E}_{pq}$ commutes with these operators. We start now with the process of defining the first-order interacting space. Here we note that the operator $\hat{E}_{ai}$ will, when it operates on a single closed-shell SD, generate two determinants, one in which an α electron has been replaced and another for the β electron. Thus, we have that an orthonormalized singly excited state is expressed as,

$$\Psi_i^a = \frac{1}{\sqrt{2}}\hat{E}_{ai}\Psi^{(0)} \ , \tag{10.92}$$

since

$$\langle\Psi_j^b|\Psi_i^a\rangle = \langle\Psi^{(0)}|\hat{E}_{jb}\hat{E}_{ai}|\Psi^{(0)}\rangle = 2\delta_{ij}\delta_{ab} \ .$$

Again, for the reasons discussed above, we need excited states that are doubly excited relative to the reference function if we are to form the first-order correction to the wave function. We now explore

the normalization using a pair of replacement operators. We define doubly excited states as

$$\Psi_{ij}^{ab} = \hat{E}_{bj}\hat{E}_{ai}\Psi^{(0)} \,, \tag{10.93}$$

for which we have the squared norm

$$\langle \Psi_{ij}^{ab} | \Psi_{ij}^{ab} \rangle = 4 + 4\delta_{ij}\delta_{ab} - 2\delta_{ij} - 2\delta_{ab} \,. \tag{10.94}$$

This is consistent with the fact that each replacement operator generates two determinants when operating on a closed-shell determinant. In the trivial case of no equal indices, $i \neq j$ and $a \neq b$, applying the two replacement operators we should get four determinants. This leads us to conclude that we need to incorporate the normalization of the excited states as we use them in our equations. Let us now explore the orthogonality within the set under a permutation of the indices. A double permutation, which translates to a change of order in the application of the replacement operators, does not cause any difference in the generated excited states, since

$$\hat{E}_{bj}\hat{E}_{ai} = \hat{E}_{ai}\hat{E}_{bj} + \delta_{aj}\hat{E}_{bj} - \delta_{bj}\hat{E}_{aj} \,, \tag{10.95}$$

where the two latter terms for obvious reasons disappear, that is $\Psi_{ij}^{ab} = \Psi_{ji}^{ba}$. However, under the condition of a single permutation of the virtual or the occupied indices, we find that the generated determinants are not identical, but they obey the relation

$$\langle \Psi_{ij}^{ba} | \Psi_{ij}^{ab} \rangle = \langle \Psi_{ji}^{ab} | \Psi_{ij}^{ab} \rangle = -2 - 2\delta_{ij}\delta_{ab} + 4\delta_{ij} + 4\delta_{ab} \,, \tag{10.96}$$

which implies that we not only have to normalize the states generated by Eq. (10.93): we have to orthonormalize them. Given the above result in the spin-specific approach we are guaranteed that the generated excited states are not redundant, except, of course, when they are identical by not properly constraining the indices in Eq. (10.93). Hence, we should not have to worry about redundancy and its elimination – this has to be considered for the case of a multi-configurational reference function. We note that some, if not all, of the non-orthogonality can be removed by forming the appropriate linear combinations as $\Psi_{ij}^{ab,\pm} = \Psi_{ij}^{ab} \pm \Psi_{ij}^{ba}$, corresponding to a singlet and triplet coupling of the virtual orbitals. This is achieved in a similar manner to how configurations of two electrons in two orbitals are linearly combined to produce singlet and triplet spin eigenstates with $M_s = 0$.

Let us, for the sake of convenience, simply index the doubly excited states generated in Eq. (10.93) with a compound index $I = I(i, j, a, b)$, such that $\Psi_{ij}^{ab} = \Psi_I$. To resolve the issue of orthonormalization we form the overlap matrix

$$S_{IJ} = \langle \Psi_I | \Psi_J \rangle \,, \tag{10.97}$$

which we subsequently diagonalize

$$\boldsymbol{s} = \boldsymbol{U}^\dagger \boldsymbol{S} \boldsymbol{U}, \tag{10.98}$$

where $\boldsymbol{s}$ is a diagonal $(L \times L)$ matrix with the eigenvalues and $\boldsymbol{U}$ is the eigenvector matrix. Since in closed-shell MPPT L can be limited to run over just the non-redundant space of double excitations, $\boldsymbol{U}$ is guaranteed to be square. However, in the general case the size of $\boldsymbol{U}$ will be $(M \times L)$, where M is the size of the original partially redundant double excitation space. However, we proceed by combining the eigenvectors with the eigenvalues according to

$$\tilde{\boldsymbol{U}} = \boldsymbol{U}\boldsymbol{s}^{-\frac{1}{2}} \,. \tag{10.99}$$

We now have a set of orthonormalized doubly excited states which are computed as

$$\tilde{\Psi}_I = \sum_J \tilde{U}_{JI}\Psi_J \,. \tag{10.100}$$

This will then render the equations for the computation of the coefficients in the first-order correction to the wave function and the second-order correction to the energy to be, starting with the former,

$$\Psi^{(1)} = \sum_I c_I \tilde{\Psi}_I \tag{10.101}$$

with the coefficients as

$$c_I = \frac{\langle \Psi^{(0)} | \hat{V} | \tilde{\Psi}_I \rangle}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b} \,, \tag{10.102}$$

where the indices of the eigenvalues of the Fock operator are resolved for the compound index I. The second-order energy correction is now computed as

$$E^{(2)} = \sum_I \frac{\langle \Psi^{(0)} | \hat{V} | \tilde{\Psi}_I \rangle^2}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b} \,. \tag{10.103}$$

Before we are finished here, let us break this down one step further and use a matrix representation all along. First we define the following matrices and vectors,

$$(\boldsymbol{H}_0)_{IJ} = \langle \Psi_I | \hat{H}_0 | \Psi_J \rangle \tag{10.104}$$

and

$$\boldsymbol{V}_I = \langle \Psi_I | V | \Psi^{(0)} \rangle \,. \tag{10.105}$$

We transform these matrices to the space of the orthonormalized excited states,

$$\tilde{\boldsymbol{H}}_0 = \tilde{\boldsymbol{U}}^\dagger \boldsymbol{H}_0 \tilde{\boldsymbol{U}} \tag{10.106}$$

and

$$\tilde{\boldsymbol{V}} = \tilde{\boldsymbol{U}}^\dagger \boldsymbol{V}. \tag{10.107}$$

The coefficients $\tilde{\boldsymbol{C}}$ in the expression for the first-order correction to the wave function, are found by solving

$$(\tilde{\boldsymbol{H}}_0 - E^{(0)}\mathbf{1})\tilde{\boldsymbol{C}} = -\tilde{\boldsymbol{V}}, \tag{10.108}$$

and the second-order correction to the energy is expressed as

$$E^{(2)} = \tilde{\boldsymbol{C}}^\dagger \tilde{\boldsymbol{V}}. \tag{10.109}$$

A final note here, in this approach we will be forced to use a two-step approach – first compute the coefficients and then the energy. With this we are now ready to discuss some particularities of the CASPT2 and some other related methods.

10.3 State-Specific Multi-Configurational Reference Perturbation Methods

After this rather lengthy introduction to perturbation theory, in this section we will, in chronological order, present the four most popular versions of multi-configurational reference perturbation theory, namely the complete active space second-order Møller–Plesset (CAS-MP2) [11], the complete active space second-order perturbation theory (CASPT2) [12, 13], the multi-reference second-order Møller–Plesset (MRMP2) [14–17] and the n-electron valence state second-order

perturbation theory (NEVPT2) [18–20]. In particular, we will devote special attention to the CASPT2 method and present it in some detail, while the three other approaches are presented more briefly and in order to mention the fundamental differences between the approaches. Briefly, the major differences between these four methods are associated with the initial selection of the reference functions $\Psi^{(0)}$ and the partitioning of the Hamiltonian into the reference Hamiltonian, $\hat{H}_0$, and the perturbation potential, V. We would also like to point out that the three latter methods CASPT2, MRMP2 and NEVPT2 have in common a *diagonalize-then-perturb* approach. That is, the CASSCF equations are effectively solved via a diagonalization-like approach and the dynamic correlation is introduced using perturbation technique. The CAS-MP2 method, on the other hand, is a hybrid somewhere in between a *diagonalize-then-perturb* and a *perturb-then-diagonalize* approach (see section 10.4 et seq.). The drawbacks and advantages of the four different approximations are discussed. In addition to these four versions of multi-configurational reference perturbation theory, a number of other flavors exist. They will not, however, be discussed in this chapter, mostly because they have not, to date, made any significant impact on standard applications. Those desiring a complete list of other such methods are recommended to consult the review paper by Lischka and co-workers [21] and specifically the subsection on "Mutireference [sic] Perturbation Theory".

10.3.1 The Generation of the Reference Hamiltonian

Before we proceed, however, let us address an issue relevant to most versions of multi-configurational reference perturbation theory beyond single determinant RSPT: What to do if (*i*) our reference function is not an eigenfunction to $\hat{H}_0$, and (*ii*) we have only exact knowledge about the reference functions, and not about the complementary functions that extend to the full Hilbert space. In this context there would be a problem to apply RSPT to known wave function models. Let us address the first problem here by constructing a new operator to solve the problem as follows. Let us assume that we have an operator, $\hat{O}$, for which our reference and complementary states are not eigenfunctions, but which we for some practical reason want to use. Then an initial operator can be constructed for the model space – the space spanned by all our reference functions – as

$$\hat{\tilde{O}} = \sum_i |\Psi_i^{(0)}\rangle\langle\Psi_i^{(0)}|\hat{O}|\Psi_i^{(0)}\rangle\langle\Psi_i^{(0)}| \,, \tag{10.110}$$

for which now all the functions $\Psi_i^{(0)}$ of the model space are eigenfunctions. This has now solved the first issue. The second problem is related to the fact that in most quantum chemistry simulations one does not have explicit access to the eigenfunctions of the complementary space. Hence, let us further assume that we can construct a complementary space to our model space, with functions that are all orthonormalized. To proceed we note that to the general expression above we are at liberty to add any complementary space we can construct, as long as these functions are mutually orthonormal. In that respect we can rewrite the expression above as

$$\hat{\tilde{O}} = \sum_{i=1}^{d} |\Psi_i^{(0)}\rangle\langle\Psi_i^{(0)}|\hat{O}|\Psi_i^{(0)}\rangle\langle\Psi_i^{(0)}| + \sum_{i=d+1}^{\infty} |\psi_i\rangle\langle\psi_i|\hat{O}|\psi_i\rangle\langle\psi_i| \,, \tag{10.111}$$

where $\Psi_i^{(0)}$ are our d reference functions, which we know exactly, and ψ_i are artificial functions with the properties $\langle\psi_i|\psi_j\rangle = \delta_{ij}$ and $\langle\Psi_i^{(0)}|\psi_j\rangle = 0$. This now solves the second problem – the lack of a complete knowledge of all eigenfunctions. The remaining obstacle is the construction of the artificial complementary space. Here a standard procedure is used, applying sets of spin-averaged electron replacement operators to our known reference wave functions – in MPPT this procedure actually leads to true eigenfunctions of the molecular Fock operator – combined with the operator

$\hat{Q} = 1 - \hat{P}$ (see Eq. 10.18) and the orthonormality conditions, to generate arbitrary complementary spaces.

10.3.2 CAS-MP2 Theory

In 1988 McDouall and co-workers published the first practical solution for extending normal single-configuration MPPT to the case of multi-configuration reference functions, in particular using a CASSCF reference function [11]. Their procedure is based on a Löwdin-type partitioning [22] of the full Hamiltonian into the subspace of the electronic configurations (Slater determinants or configuration state functions), and the complementary space. The resulting equations are subsequently solved using simple MPPT. Let us outline, in some detail, the approach here. We define a model space that includes all the CASSCF states. In this aspect we have that the projection operators $\hat{P}$ and $\hat{Q}$ play the same role as in the section on the multi-state case (see section 10.4). In this formalism we divide the Hamiltonian by blocks into P and Q spaces. In particular we note that $\hat{P} = \sum_i |\Psi_i^{(0)}\rangle\langle\Psi_i^{(0)}|$, where $\Psi_i^{(0)}$ are solutions to the CASSCF equations. We then introduce a transformation, $\boldsymbol{U}$, which transforms the Hamiltonian into a representation in which the P and Q spaces are decoupled, that is

$$\begin{pmatrix} \boldsymbol{H}^{PP} & \boldsymbol{H}^{PQ} \\ \boldsymbol{H}^{QP} & \boldsymbol{H}^{QQ} \end{pmatrix} \boldsymbol{U} = \boldsymbol{U} \begin{pmatrix} \boldsymbol{H}^{\text{eff}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{W}^{*} \end{pmatrix} . \tag{10.112}$$

In this matrix representation $\boldsymbol{H}^{PP} = \boldsymbol{PHP}$ and so on. For the model space we have the normal CASSCF solutions to the time-independent Schrödinger equation,

$$\hat{P}\hat{H}\hat{P}\Psi_i^{(0)} = E_i^{(0)}\Psi_i^{(0)}. \tag{10.113}$$

The transformation matrix, $\boldsymbol{U}$, will have diagonal blocks in which the transformation is arbitrary, for the sake of convenience we set the matrix to have the form

$$\boldsymbol{U} \equiv \begin{pmatrix} \boldsymbol{1} & -\boldsymbol{C}^{\text{T}} \\ \boldsymbol{C} & \boldsymbol{1} \end{pmatrix} , \tag{10.114}$$

where $\boldsymbol{C}$ is a set of column vectors, one for each function in the model space. This now generates two working equations,

$$(\boldsymbol{H}^{\text{eff}})_{ij} = (\boldsymbol{H}^{PP})_{ij} + (\boldsymbol{H}^{QP})_i^{\text{T}}\boldsymbol{C}_j \tag{10.115}$$

and

$$(\boldsymbol{H}^{QP})_i + \boldsymbol{H}^{QQ}\boldsymbol{C}_i = \sum_j (\boldsymbol{H}^{\text{eff}})_{ji}\boldsymbol{C}_j , \tag{10.116}$$

where two subindices indicate a matrix element, and one subindex indicates a whole column. The total energy for state i is then approximated as

$$E_i \approx \langle\Psi_i^{(0)}|\hat{H}^{\text{eff}}|\Psi_i^{(0)}\rangle . \tag{10.117}$$

McDouall and co-workers point out that this approach – a hybrid approach between a *diagonalize-then-perturb* and a *perturb-then-diagonalize* scheme – comes with two advantages. First, single-excitation contributions are zero, and second, the off-diagonal correction of Eq. (10.115) is expected to be small and can be ignored. The energy is now expressed as

$$E_i \approx E_i^{(0)} + \langle\Psi_i^{(0)}|(\hat{H}^{QP})_i^{\text{T}}\hat{C}_i|\Psi_i^{(0)}\rangle , \tag{10.118}$$

where $\hat{X}_i$ is the operator representation of the ith column of matrix $\boldsymbol{X}$. What remains now is to compute $\boldsymbol{C}$ using MPPT theory. The authors note that one can approximate Eq. (10.116) by restricting the summation on the RHS to only include the same $\boldsymbol{C}$-vector as on the LHS – i.e., taking only $j = i$, an approximation which affects the results only at the third order of the perturbation. This leaves us the following equation,

$$\boldsymbol{C}_i = -(\boldsymbol{H}^{\mathrm{QQ}} - (\boldsymbol{H}^{\mathrm{eff}})_{ii}\mathbf{1})^{-1}(\boldsymbol{H}^{QP})_i \ . \tag{10.119}$$

We can now proceed solving this by applying the follow procedure. First, the inverse problem in Eq. (10.119) is approached with the use of a power series expansion. Second, sums over states are replaced with sums over orbitals. Third and finally, the diagonal elements of the matrix to invert are replaced with orbital energy differences from the diagonal elements of some reference one-electron Hamiltonian $\hat{h}_i$. To be explicit, the authors used the power series expansion

$$(\boldsymbol{A} - \boldsymbol{B})^{-1} = \boldsymbol{A}^{-1} + \boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{A}^{-1} + \dots , \tag{10.120}$$

where $\boldsymbol{A}$ is the matrix of the diagonal elements of $\boldsymbol{H}^{\mathrm{QQ}} - (\boldsymbol{H}^{\mathrm{eff}})_{ii}\mathbf{1}$. Moreover, for the reference one-electron Hamiltonian the authors used different $\hat{h}_i$ for each reference state i. This makes sense since in some reference state (read determinant or possibly configuration state function) an active orbital might be occupied while in another state it is empty. Using a unified $\hat{h}$ operator under these conditions would render the perturbation expansion to diverge. However, the rules for the definition of the $\hat{h}_i$ are to some extent *ad hoc* and sensitive to the orbital representation, McDouall and co-workers have selected to use a localization of the active orbitals leading to an orthogonal valence-bond expansion. As for the performance with respect to computer time the authors state that the method will cost n times the cost of a single-reference MP2, where n is the number of functions in the reference space. This fact means that the approach is not often used. Furthermore, the literature lacks any benchmarks comparing this approach to any of the other MRPT methods presented in this section. From a theoretical ground there is no reason for the CAS-MP2 approach to be worse than the methods that will follow, rather it might actually have a slight edge due to the use of a hybrid approach. However, the poor scaling with the size of the CI expansion will, in general, render the approach not useful.

10.3.3 CASPT2 Theory

Here we present the mathematical equations which govern the complete active space second-order perturbation theory (CASPT2) method, a second-order perturbation method.[1] More specifically, we will present the state-specific CASPT2 method (SS-CASPT2), in which each state is treated independently. In this approach we note that, if we have a perturbation series which is convergent, each of the states we consider will, as we go towards infinite order, get the energy and wave function correct. We will be able to compute the ground and any excited states which are spanned by the active space of the reference CASSCF wave function. In general this will work very well if the original CASSCF states are good qualitative representatives of the final states. However, convergence will be slow if the reference states are poor and the successive iterations over the order of the perturbational strength parameter are desperately needed for the corrected wave function to mirror the exact wave function of the state under study in a qualitative way. Examples of this can be numerous, for instance, the lack of dynamic electron correlation in the CASSCF model can cause

1 A review of the state-specific CASPT2 method was presented by Andersson and Roos in 1995 [23]. This section is an extension of that review and we recommend students to consult that material in parallel.

the reference functions to artificially mix states beyond the one they are trying to represent to be used in a perturbational approach. We will present below approaches that will fix these problems, but more on that later.

The CASPT2 method is based on (*i*) a multi-configurational reference wave function, the CASSCF wave function, as described in detail in Chapter 6, and (*ii*) a partitioning of the Hamiltonian, such that it reduces down to Møller–Plesset partitioning, that is, the definition of $\hat{H}_0$ is the same – a one-electron operator – if the CASSCF wave function is reduced down to a single-determinant wave function. The selection of $\hat{H}_0$ is for reasons of simplicity required to be a one-electron operator, alternative but similar partitionings have been introduced [24] but have over the years been abandoned due to suboptimal performance. As we proceed we face two issues. First, we need a generalization of the Fock operator to wave functions of arbitrary spin and number of open shells. Second, we need to formulate a $\hat{H}_0$ for which the CASSCF wave function is an eigenfunction.

10.3.3.1 The Partitioning of the Hamiltonian

Let us start with the mathematical description of the SS-CASPT2 method, as it is applied to a general CASSCF state. The generalized Fock operator, the $\hat{H}_0$ operator of the SS-CASPT2 method, is defined as

$$\hat{F}(D) = \sum_{pq} f_{pq}(D)\hat{E}_{pq} \ , \tag{10.121}$$

where D is the CASSCF one-particle reduced density matrix of the state of interest, represented by the state $\Psi^{(0)}$. The scalar value f_{pq} is computed as

$$f_{pq}(D) = \frac{1}{2}\sum_{\sigma=\alpha}^{\beta}\langle\Psi^{(0)}|[\hat{a}^{\dagger}_{p_\sigma},[\hat{a}_{q_\sigma},\hat{H}]]|\Psi^{(0)}\rangle \tag{10.122}$$

$$= h_{pq} + \sum_{rs} D_{rs}\left[(pq|rs) - \frac{1}{2}(pr|qs)\right] \ , \tag{10.123}$$

which in the case of the standard closed-shell single-determinant case will boil down to the conventional matrix elements of the molecular orbital Fock operator. In passing we note that the matrix $\boldsymbol{f}$ formed by the f_{pq} values can be broken down into nine blocks based on the division of the orbitals into inactive, active and virtual orbitals,

orbital class	inactive	active	virtual
inactive	ϵ_i	f_{it}	0
active	f_{ti}	ϵ_t	f_{ta}
virtual	0	f_{at}	ϵ_a

The diagonal block of this matrix (green boxes above) can be diagonalized independently resulting in four remaining non-zero blocks (red boxes) (the inactive–active, active–inactive, active–virtual, and virtual–active blocks) and two zero blocks (the inactive–virtual and virtual–inactive blocks). In the initial implementation of CASPT2 the Fock operator was approximated to be diagonal to simplify the calculation and compute the second-order energy and coefficients of the first-order wave function simultaneously [12]. However, this approximation, called D-CASPT2, proved to be too severe and later an iterative implementation was presented in which the non-diagonal nature

of the Fock matrix is kept [13], the N-CASPT2 methods. This prefix has now been dropped, the D-CASPT2 results are very poor and all CASPT2 results now refer to the use of the full Fock matrix.

10.3.3.2 The First-Order Interacting Space

We have now solved our first problem – we have a $\hat{H}_0$ operator, however, in contrast to MP theory our CASSCF reference wave functions are not eigenfunctions to the generalized Fock operator. In this respect we will have to follow the recipe described a couple of sections above, that is we introduce the identity as defined by the projection operators

$$1 = |\Psi^{(0)}\rangle\langle\Psi^{(0)}| + \hat{P}_{\mathrm{K}} + \hat{P}_{\mathrm{SD}} + \hat{P}_{\mathrm{TQ}} + \dots , \tag{10.124}$$

in which the projection operators $\hat{P}_{\mathrm{K}}$, $\hat{P}_{\mathrm{SD}}$, etc., are defined as

$$\hat{P}_X = \sum_{i \in X} |\Psi_i\rangle\langle\Psi_i| ,$$

where Ψ_i is a orthonormal state of the X space ($X \in \{\mathrm{K}, \mathrm{SD}, \mathrm{TQ}, \dots\}$). More specifically, $\hat{P}_{\mathrm{K}}$ projects into the complementary space, wave functions corresponding to all other possible wave functions that are a solution in the CAS space, but are not in the reference space. $\hat{P}_{\mathrm{SD}}$ corresponds to states which are single or double replacements (SD)[2], relative to the reference function, generated as $\hat{E}_{pq}\hat{E}_{rs}|\Psi_i^{(0)}\rangle$. Note that, as in the case of MPPT, single replacements can be safely ignored, since for a variationally optimized CASSCF wave function their interaction with $\Psi^{(0)}$ is zero (Brillouin–Levy–Berthier theorem). We selected this approach before for a reason. Let us visit the SD space in a bit more detail, since if we are only to compute second-order corrected energies we need only worry about the design and structure of this space. In the SD space the electron replacements (see below) are complete or partial replacements out of the inactive–active space into the active–virtual space. This includes the so-called semi-internal replacements, involving the coupling of an internal replacement (inactive–active replacement) with an external replacement (inactive–virtual or active–virtual promotion). To present all the SD space, it is divided into eight classes of operators acting on the reference wave function ($V_{\mathrm{A}}, \dots, V_{\mathrm{H}}$), they are presented as:

$$\begin{aligned}
&\text{internal} \begin{cases} V_A\,, & k = \;\; 1 : \hat{E}_{ti}\hat{E}_{uv} \\ V_B\,, & k = \;\; 2 : \hat{E}_{ti}\hat{E}_{uj} \end{cases} \\
&\text{semi-internal} \begin{cases} V_C\,, & k = -1 : \hat{E}_{at}\hat{E}_{uv} \\ V_D\,, & k = \;\; 0 : \hat{E}_{ai}\hat{E}_{tu}\,,\ \hat{E}_{ti}\hat{E}_{au} \\ V_E\,, & k = \;\; 1 : \hat{E}_{ti}\hat{E}_{aj} \end{cases} \\
&\text{external} \begin{cases} V_F\,, & k = -2 : \hat{E}_{at}\hat{E}_{bu} \\ V_G\,, & k = -1 : \hat{E}_{ai}\hat{E}_{bt} \\ V_H\,, & k = \;\; 0 : \hat{E}_{ai}\hat{E}_{bj}\,, \end{cases}
\end{aligned}$$

where i and j are inactive orbital, a and b are virtual orbital, and t, u, and v are active orbital indices of the CASSCF wave function. The classification as internal, semi-internal or external corresponds to the number of virtual indices, 0, 1 or 2, respectively. The index k denotes the change of number of electrons in the active space under the application of the replacement operator – this is used in the NEVPT2 terminology to divide the first-order interacting space into subclasses

2 It is a confusing fact that both Slater determinants and the single and double replacement space of the first-order interacting space in the literature are abbreviated SD. However, we hope that it will be clear from the context what the abbreviation refers to.

(see section 10.3.5 below). We note that the last case corresponds to the normal set of parameters found in conventional MP2 theory and they are the ones that are the largest set. As in the case of MPPT, non-orthogonalities in the first-order interacting space can be removed by taking linear combinations of operators, as in $\hat{E}_{ai}\hat{E}_{bj} \pm \hat{E}_{bi}\hat{E}_{aj}$.

Despite the block nature of $\hat{F}(D)$, we will not get a block-diagonal Hamiltonian – this would translate to the correction vectors spanning the full Hilbert space. To overcome this we construct a projected $\hat{H}_0$ operator as

$$\begin{aligned}\hat{H}_0 = &|\Psi^{(0)}\rangle\langle\Psi^{(0)}|\hat{F}(D)|\Psi^{(0)}\rangle\langle\Psi^{(0)}| \\ &+ \hat{P}_{\mathrm{K}}\hat{F}(D)\hat{P}_{\mathrm{K}} + \hat{P}_{\mathrm{SD}}\hat{F}(D)\hat{P}_{\mathrm{SD}} \\ &+ \hat{P}_{\mathrm{TQ}}\hat{F}(D)\hat{P}_{\mathrm{TQ}} + \dots .\end{aligned} \tag{10.125}$$

Thus, this $\hat{H}_0$ operator will be block-diagonal as follows

space	0	K	SD	TQ	⋯
0	$\boldsymbol{H}_0^{(0,0)}$	**0**	**0**	**0**	**0**
K	**0**	$\boldsymbol{H}_0^{(\mathrm{K},\mathrm{K})}$	**0**	**0**	**0**
SD	**0**	**0**	$\boldsymbol{H}_0^{(\mathrm{SD},\mathrm{SD})}$	**0**	**0**
TQ	**0**	**0**	**0**	$\boldsymbol{H}_0^{(\mathrm{TQ},\mathrm{TQ})}$	**0**
⋮	**0**	**0**	**0**	**0**	⋱

which will assure us that the first-order correction to the wave function is completely defined by the SD subspace.

Proceeding with a brief analysis of the SD space we first note that we will have a set of excited states which are neither normalized nor orthogonal. Even worse, they are not necessarily linearly independent. Hence, orthonormalization and elimination will have to be included in the process in which the overlap matrix $\boldsymbol{S}_{IJ}$ in Eq. (10.97) is processed (block by block) to generate a non-redundant set of orthonormalized excited states. This is now used to make the appropriate transformations to generate the matrix $\tilde{\boldsymbol{H}}_0$ and the vector $\tilde{\boldsymbol{V}}$. All this is used in the equation to derive the coefficients of the first-order corrections to the wave function (see Eq. 10.108). Second, different to standard MPPT, with the use of a single closed-shell determinant as a reference wave function, the matrix of the unperturbed Hamiltonian will not be trivially diagonal, while still block-diagonal. A brute force approach would be to diagonalize the matrix on the LHS, however, that would limit the size of the calculations we could do – the space of the doubly excited states would be too large for the explicit storage of the $\tilde{\boldsymbol{H}}_0$ matrix. Rather the Equation will have to be solved in an iterative fashion using a preconditioned conjugate gradient approach computing vectors (known as σ vectors) of the form

$$\sigma = (\tilde{\boldsymbol{H}}_0 - E^{(0)}\mathbf{1})\boldsymbol{C}^{(1)} , \tag{10.126}$$

where $\boldsymbol{C}^{(1)}$ is a trial vector of the coefficients of the excited states in the SD subspace. In this context we will evaluate the matrix on the fly as we make the contraction (product) with the trial vector. That is, the matrix $\tilde{\boldsymbol{H}}_0 - E^{(0)}\mathbf{1}$ is not stored explicitly. Rather, as an element of the σ vector is generated only those parts of the $\tilde{\boldsymbol{H}}_0 - E^{(0)}\mathbf{1}$ matrix which are required are generated on the fly.

Moreover, as the projection operator, $\hat{P}$, is extended further to sets of two further replacements – the TQ manifold, and so on – we achieve a continued block diagonalization of $\hat{H}_0$ over the blocks of the reference function, the complementary CAS space, the SD space, the TQ space, etc.

We are now at the point at which we need to form the matrix elements (I, J are compound indices for the states in the SD space)

$$\begin{aligned}(\boldsymbol{H}_0)_{IJ} &= \langle \Psi_I | \hat{H}_0 | \Psi_J \rangle \ , \\ \boldsymbol{S}_{IJ} &= \langle \Psi_I | \Psi_J \rangle \ , \text{and} \\ \boldsymbol{V}_I &= \langle \Psi_I | \hat{V} | \Psi^{(0)} \rangle \ ,\end{aligned}$$

where

$$\hat{V} = \sum_{pqrs} \hat{e}_{pqrs}(pq|rs) - \hat{E}_{pq} \sum_{rs} D_{rs} \left[(pq|rs) - \frac{1}{2}(pr|qs) \right] \ . \tag{10.127}$$

We will use these matrices and vectors to (*i*) eliminate linear dependence, (*ii*) introduce a non-redundant orthonormal space of states in the SD space, (*iii*) commence the iterative procedure to compute the coefficients of the first-order correction to the wave function, and (*iv*) compute $E^{(2)}$.

Before we continue let us examine, in some detail, the matrix elements which have to be computed and what this is associated with (a complete list of all required equations for the compilation of the CASPT2 energy are collected in the Appendix). We have that the wave functions Ψ_I and Ψ_J contain the product of two one-electron replacement operators, and so does $\hat{V}$, while $\hat{H}_0$ is a one-electron operator and contains single replacement operators. Therefore, the worst case could be expected to be the computation of matrix elements for $\hat{H}_0$. For example, for one of the external cases (V_{H}), we need to evaluate expressions such as

$$\sum_{pq} \langle \Psi_i^{(0)} | \hat{E}_{kc} \hat{E}_{ld} \hat{E}_{pq} \hat{E}_{ai} \hat{E}_{bj} | \Psi_j^{(0)} \rangle \ . \tag{10.128}$$

This and other expressions seem to indicate that *n*-particle density matrices up to five-particle matrices would be needed, since there are five replacement operators. However, the general reduction (for a complete list of the equations up to the second-order correction of the energy see the Appendix) will be such that the Hamiltonian will eliminate the four virtual indices, that is, a CASPT2 implementation needs, in addition to the one- and two-particle density matrices, three- and partially contracted four-particle density matrices (effectively of the size of a three-particle density matrix). The actual storage and computation cost for all the *n*-particle density matrices (1-, 2-, 3- and partially contracted 4-particle density matrices) will scale in the range of m^6–m^8, where m is the number of active orbitals. This seems to be tractable right now but, as will be seen later in section 10.3.6, the density matrix renormalization group approach to CASSCF calculations will facilitate large active spaces, at which point, when m can easily be of the order 10^2, the computation and handling of the density matrices can become a serious bottleneck.

We are almost done. However, we should not forget to address the issue of size-extensivity of the CASPT2 method. The CASSCF and MP2 methods (the latter at least in the case of closed-shell systems) are size-extensive. Does this hold for the CASPT2 method too? Helgaker and co-workers [3] show that since $\hat{H}_0$ is not additively separable (due to the projection operators present in the $\hat{H}_0$ definition), the CASPT2 model is not a size-extensive perturbation model. Further, they note that as the reference wave function increases the multi-configurational character, the extent of the non-size-extensivity increases.

10.3.3.3 Other Active Space References

The CASPT2 method, as should be clear from the preceding and following sections, is not free of issues. To the just mentioned non-size-extensivity, we can add the possibility of intruder states (10.1.2), the overstabilization of open-shell states (10.3.3.5) and the non-orthogonality of the first-order wave functions (10.5). It should be noted that these problems are not specific to CASPT2, and other multi-configurational reference perturbation methods can also show some or all of them to some extent. There is, however, a simple – and not particularly useful – recipe to solve or mitigate these problems: increase the active space size. From a purely theoretical point of view, by increasing the active space size we approach the full CI solution, at which point our reference CASSCF functions *are* the exact solutions, the zeroth-order Hamiltonian $\hat{H}_0$ is the exact Hamiltonian and the perturbation V vanishes. It "solves" the problems by making the PT treatment superfluous. From a practical standpoint, the exponential increase in the number of SDs with the size of the active space makes the required CASSCF calculations intractable for anything beyond about 20 orbitals, not to mention the PT2 step.

As described in Chapter 6, other approaches have been proposed to allow extending an active space beyond what is practically possible with CASSCF, while keeping the number of SDs manageable. Two of these approaches are the restricted active space SCF (RASSCF) and generalized active space SCF (GASSCF) methods. How can we apply a perturbative treatment to reference functions based on RASSCF or GASSCF, instead of CASSCF?

The corresponding extensions of CASPT2, RASPT2 [25] and GASPT2 [26], have been published relatively recently. In the case of RASPT2, the matrix representation of the Fock operator can be easily obtained from the one in CASPT2, noting that now the active space is split into three different subspaces:

			active			
	orbital class	inactive	RAS1	RAS2	RAS3	virtual
	inactive	ϵ_i	f_{it}			0
active	RAS1	f_{ti}	ϵ_t	$f_{tt'}$	$f_{tt'}$	f_{ta}
active	RAS2		$f_{tt'}$	ϵ_t	$f_{tt'}$	
active	RAS3		$f_{tt'}$	$f_{tt'}$	ϵ_t	
	virtual	0	f_{at}			ϵ_a

where we see that the active–active block cannot be fully diagonalized, because the different RAS spaces cannot be mixed without altering the reference wave function. The first-order interacting space for a RASPT2 wave function contains all the eight classes considered for CASPT2 plus a fully active class given by $\hat{E}_{tu}\hat{E}_{vx}$, where all indices refer to active orbitals. This class is absent in CASPT2 because all possible excitations within the active space are already included in the reference or complementary space, but in RASPT2 some of these excitations would violate the RASSCF restrictions and should therefore be included in the first-order interacting space. In GASPT2, for every pair of not fully connected GAS spaces (where not all possible interspace excitations are allowed), there exists a fully active class as in RASPT2.

The implementation of a rigorous PT2 treatment of such a structure poses some challenges. First, including active–active excitations would require sorting out those that still fulfill the RAS or GAS

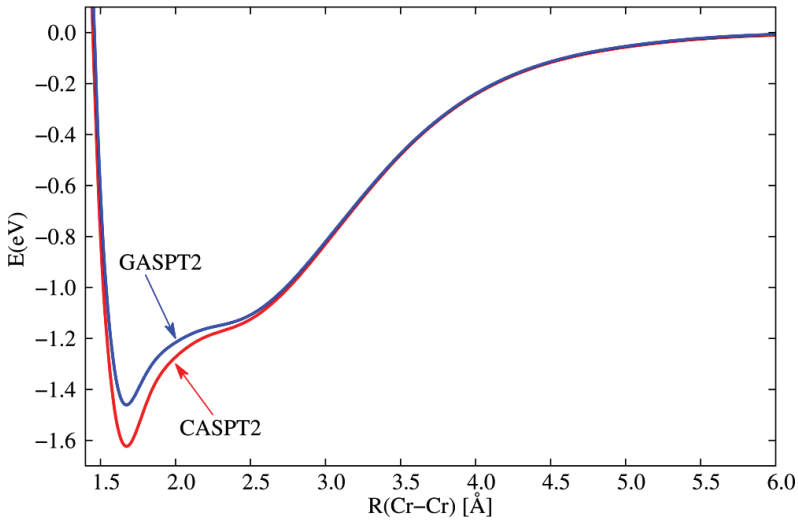


Figure 10.3 CASPT2(12,12) and GASPT2(12,12) potential energy curves for Cr_2. Reprinted with permission from ref. [26]. Copyright 2016, American Chemical Society.

restrictions from those that do not – the former should be removed, the latter should stay in the first-order interacting space. Furthermore, the computation of the matrix elements over $\hat{H}_0$ would require four-particle density matrices, which would be impractical for larger active spaces, and it is precisely for the use of large active spaces that we would like to use these methods. Therefore, in the current implementations of both RASPT2 and GASPT2, an approximation is employed in removing the active–active excitations from the first-order interacting space. The effect of this approximation is a degradation in the accuracy of the perturbative treatment, but it is expected that the possibility of using larger active spaces for the reference functions will compensate for it. As an example, fig. 10.3 shows a comparison of CASPT2 and GASPT2 curves for the difficult case of Cr_2, where a CASPT2 calculation is still affordable. The GASPT2 calculation was done by splitting the active space into six separate GAS spaces of two orbitals (bonding and antibonding) with two electrons, with no interspace excitations allowed. The difference between the two curves around the equilibrium geometry is attributed to the missing interspace excitations in the GASSCF treatment, which GASPT2 cannot recover due to the approximation employed.

10.3.3.4 Benchmark Results

A few words on the quality of the CASPT2 results are opportune at this point. The very first benchmark of the CASPT2 method was conducted by Andersson and Roos [27]. In this benchmark, molecular structures and binding energies were computed for some 27 molecules containing first- and second-row atoms (H–Ne). The results indicate an accuracy in bond distances and bond angles of 0.01 Å and 0–2 degrees, respectively. Furthermore, it is found that atomization energies are underestimated by 3–6 kcal mol^{-1} times the number of unpaired electrons created in the atomization process – the "correction" of this systematic error led to the so-called IPEA-shift technique, which is discussed below. Finally, the error in isogyric reactions – reactions which preserve the number of unpaired electrons – was in the range of −2.5–1 kcal mol^{-1}. In 2003 Guner and co-authors [28] reported benchmark results for 11 pericyclic reactions of unsaturated hydrocarbons. In particular they studied properties such as transition state structures, reaction enthalpies and entropies, and activation enthalpies and entropies. They reported, for example, mean absolute

deviations (MAD) for the reaction and activation enthalpies of 1.6 kcal mol^{-1} and 1.7 kcal mol^{-1}, respectively, at the CASPT2/6-31G*//CASSCF[3] level of approximation. Schreiber and co-workers [29] performed a comparative study of singlet (121 transitions) and triplet (71 transitions) vertical excitation energies computed at the CASPT2 and the CC3 levels. They reported MADs of 0.08 eV and 0.20 eV for the triplet and singlet transitions, respectively. These results, especially the organic spectroscopic data, were summarized by Serrano-Andrés in a special issue of IJQC [30] in honor of Björn O. Roos. In the same issue Pierloot [31] and Gagliardi [32] report on the state of the art of CASPT2 studies of transition metal compounds and actinide chemistry, respectively. Recently Budzák and collaborators established a database of ground-state and excited-state structures from some 35 molecules. Again, it was demonstrated that the CASPT2 method yields deviations of only 0.01 Å in any bond type when compared to the "gold standard" coupled-cluster methods [33].

10.3.3.5 IPEA Shift

As mentioned above, Andersson and Roos reported large deviations in atomization energies – up to 20 kcal mol^{-1}. Initially this caused much alarm, however, after some analysis it was established that this was a systematic error depending on the change of the number of unpaired electrons in the process under study. That is, an imbalance in the treatment of the reference Hamiltonian, $\hat{H}_0$, of closed-shell electrons with respect to unpaired electrons. In general this leads to too low bond and excitation energies. To remedy this artifact, Andersson introduced corrections to the Fock matrix, named G_1, G_2, and G_3 [24]. Although these corrections could relieve the problem in some cases, it was limited to situations where the energy gap between active and virtual orbitals is large. Ghigo and co-workers [34], however, suggested an alternative shift which seemed to reduce the systematic error significantly. In their analysis they identified that the systematic error could easily be resolved using a two-electron term, à la Dyall [35] (see section 10.3.5 on NEVPT2, where it is described in some detail), in the zeroth-order Hamiltonian. This would, however, lead to an increase in the computational expense, and a modification should preferably fit within the scope of a one-electron-operator-based formalism. Ghigo et al. further noted that for closed-shell orbitals the negative of the diagonal value of the Fock matrix corresponds to the ionization potential (IP), while for an empty orbital the value corresponds to an electron affinity (EA). This is trivially the case for the inactive and the virtual spaces, respectively, in the CASSCF model. For the active space, however, the situation is not that simple and the nature of the diagonal value is closely associated with the occupation number. To proceed, let us reformulate Eq. (10.123) as an interpolation between the two cases that the active orbital is doubly occupied (−IP) or empty (−EA),

$$f_{pp} = -\frac{1}{2}(D_{pp}(\mathrm{IP})_p + (2 - D_{pp})(\mathrm{EA})_p) \; . \tag{10.129}$$

Here we have an expression which is correct for $D_{pp} = 0$, 1, or 2 and in general mimics the behavior of the diagonal elements of the Fock matrix for the active orbitals. In particular this expression highlights that we will have a problem when an electron is attached to or removed from a singly occupied orbital. As a result, the denominators of the coefficients in the first-order correction of the wave function will be too small, the coefficients will be too large and thus contribute too much to the second-order correction to the energy. This is what had been empirically found – the perturbation energy is overestimated for open shells. Additionally, this will be yet another source for the manifestation of intruder states.

We would like to have a formalism such that we would use an electron affinity if we excite into an orbital and an ionization potential if we remove an electron from it. Ghigo and co-workers suggest

3 This acronym indicates that the study was conducted with a 6-31G* basis set, structures optimized at the CASSCF level, and finally the energies are computed with the CASPT2 method.

that this can be achieved by the introduction of a shift, the IPEA shift. In the case we move an electron into an orbital p, we would like to shift Eq. (10.129) by the value

$$\sigma^{\mathrm{EA}} = \frac{1}{2} D_{pp}((\mathrm{IP})_p - (\mathrm{EA})_p) \,, \tag{10.130}$$

while in the case of the removal of an electron from the same orbital, we would like to shift by the value

$$\sigma^{\mathrm{IP}} = -\frac{1}{2}(2 - D_{pp})((\mathrm{IP})_p - (\mathrm{EA})_p) \,. \tag{10.131}$$

This would be completely in line with the possibility of implementing a state-specific shift (see Eq. 10.65), where in this case the summation is limited to the states in the first-order interacting space and the shift is specific to each of the classes of states in this space. Subsequently Ghigo and co-workers took a leap of faith – considering that the definition of and the value of $(\mathrm{IP})_p - (\mathrm{EA})_p$ is not well defined nor is it in general known – they suggested the use of a single parameter σ, to be determined from accurate calculations or experimental results. They conclude their study by demonstrating, using a σ value of 0.25 E_h, that the computed dissociation energy errors for some 49 molecule can be reduced from more than 0.2 eV to less than 0.1 eV (see fig. 10.4). This value of the shift, 0.25 E_h, has since then been the default value in the MOLCAS quantum chemistry program package. [36] It should again be pointed out that this shift factor will be applied differently for each of the eight different classes of the first-order interacting space (see Eq. 10.65, and read refs. [37] and [38] for a detailed description of the implementation of the IPEA shift).

The IPEA-shift technique has, over the years, been an origin of irritation and many have criticized the approach. Some claim the shift makes the CASPT2 method semi-empirical, others state that the shift should be significantly larger, etc. [37, 39, 40]. To resolve the degree of "correctness" and validity of the IPEA shift, Zobel and co-workers [38] recently reported results from (*i*) an extensive study of literature of CASPT2 excitation energies, (*ii*) full configuration interaction benchmark results compared with CASPT2 results, and (*iii*) extensive studies using different values of σ. From the results Zobel and co-authors observed that the correction scales with the amount of dynamic correlation and is a function of the size of the basis set (see fig. 10.5). Thus they conclude that the notion of an universal IPEA shift does not have any support and that the practice should be abandoned.

10.3.4 MRMP2 Theory

About two years after the proposal and implementation of the CASPT2 theory, Hirao suggested a variation of the theory named multi-reference second-order Møller–Plesset perturbation theory (MRMP2) [14–17]. This proposal is similar in many respects to the CASPT2 theory and the differences will be briefly presented below. The major difference lies in the design of the $\hat{H}_0$ operator and the definition of the first-order interacting space.

10.3.4.1 The Partitioning of the Hamiltonian

The zeroth-order Hamiltonian is trivially defined as

$$\hat{H}_0 = \sum_i E_i^{(0)} |\Psi_i^{(0)}\rangle\langle\Psi_i^{(0)}| \,, \tag{10.132}$$

where $\Psi_i^{(0)}$ are CASSCF wave functions and $E_i^{(0)}$ are the reference energies defined as follows. First, the reference energy is defined as a sum over orbital-energy-like terms, generalized orbital energies,

$$E_i^{(0)} = \sum_p \epsilon_{pp} \langle\Psi_i^{(0)}|\hat{E}_{pp}|\Psi_i^{(0)}\rangle \,, \tag{10.133}$$

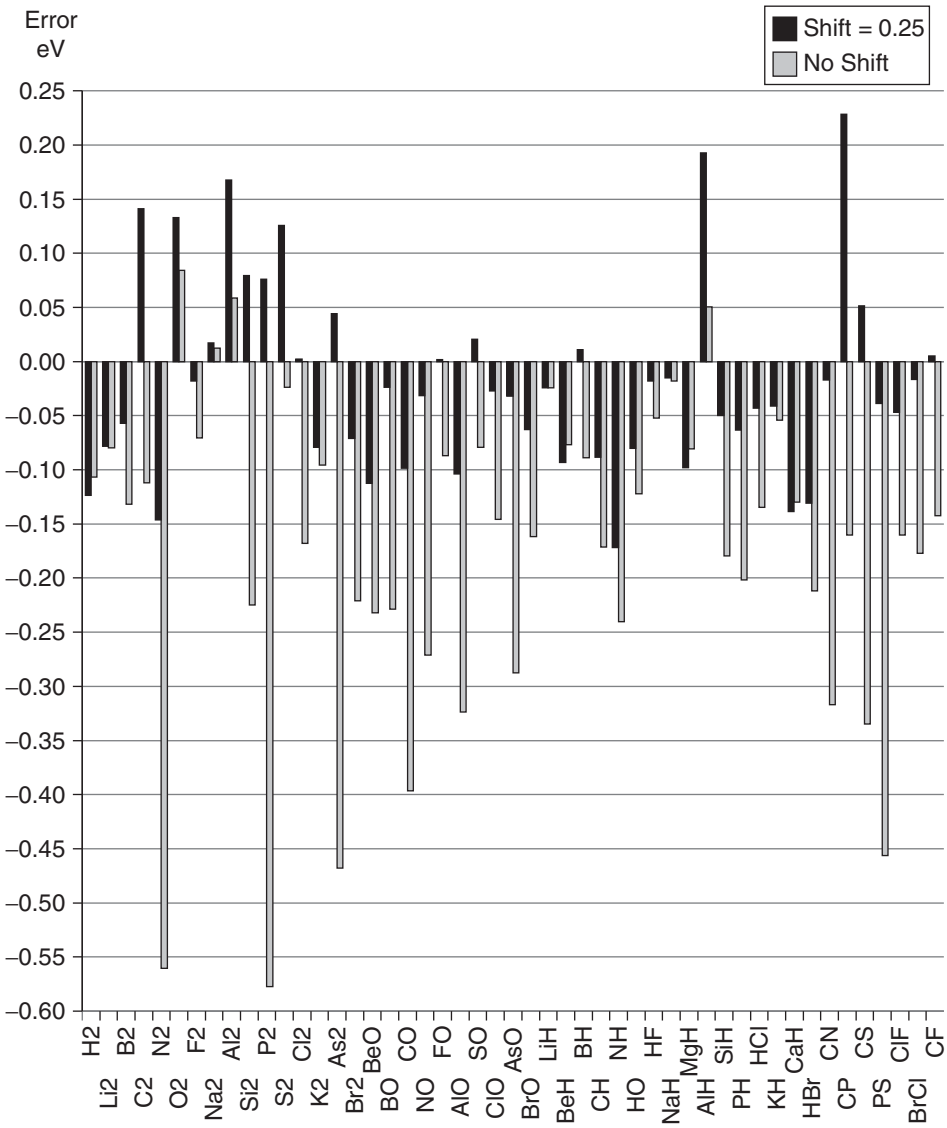


Figure 10.4 Errors in the dissociation energy for 49 diatomic molecules obtained with the IPEA shift parameter equal to 0.25 E_h and with no IPEA shift. Reprinted with permission from ref. [34]. Copyright 2004, Elsevier.

where $\hat{E}_{pp}$ is the number operator (it returns the occupation number for the pth orbital). It is important to note that these matrix elements are given in the orbital representation that diagonalizes the diagonal blocks of the Fock matrix. In this orbital representation the generalized orbital energies are computed as

$$\epsilon_{pq} = h_{pq} + \sum_{rs} D_{rs} \left[(pq|rs) - \frac{1}{2}(pr|qs) \right] . \tag{10.134}$$

10.3.4.2 The First-Order Interacting Space

While the CASPT2 and NEVPT2 methods (see section 10.3.5 below) define the first-order interacting space by the use of replacement operators applied to the contracted CASSCF wave function(s),

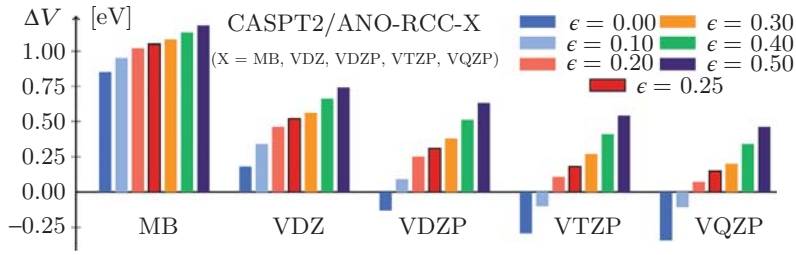


Figure 10.5 Mean signed errors (MSEE) in eV of CASPT2 vertical excitation energies compared to experimental reference data for different ANO-RCC basis sets and IPEA shift values ϵ (in E_h) Reprinted from ref. [38], published by The Royal Society of Chemistry, distributed under Attribution 3.0 Unported (CC BY 3.0) license.

the MRMP2 method defines the space by collecting all SDs that correspond to a single or double replacement relative to the SDs that span the CASSCF CI space. In that respect this space is larger than the corresponding space in, for example, the CASPT2 approach, however it comes with the benefit that the representation provides a diagonal representation of $\hat{H}_0$. This space has, of course, to be subjected to elimination of linear dependences.

We conclude this section by noting that the MRMP2 method is not used extensively and there is a lack of any significant amount of benchmarking in the literature, compared to other MRPT methods.

10.3.5 NEVPT2 Theory

The *n*-electron valence state for multi-reference perturbation theory (NEVPT2) was introduced some 10 years after the CASPT2 method [18]. At the time three different versions of the method were introduced: the totally uncontracted, the partially contracted (PC), and the strongly contracted (SC). Different to CASPT2, NEVPT2 uses an $\hat{H}_0$ operator that includes two-electron terms, and the method is attributed to avoid the problems with intruder states, which frequently cause issues in CASPT2 studies. Thus, two of the major advantages that the method claims are (*i*) exempt from the intruder state problem and (*ii*) the method is size consistent.

The totally uncontracted method entails a high computational cost and has not been used in practical applications (for some recent developments, see the end of section 10.3.6). The two latter versions, the PC-NEVPT2 and the SC-NEVPT2 versions, however, are in common use. Hence, we will discuss these methods briefly.

10.3.5.1 The Partitioning of the Hamiltonian

Below, we will describe the $\hat{H}_0$, however, in the case of NEVPT2 this requires the generation of the functions in the first-order interacting space – also known as *perturbers*, since they will perturb the wave function as we include corrections through the perturbation expansion. We will, at this point, assume that these perturbers are generated and denoted, $\Psi_l^{\{k\}}$ and $\Psi_{l\mu}^{\{k\}}$, for the strongly and the partially contracted NEVPT2 versions, respectively. Here k and l index the class and subclass, respectively, of the first-order interacting functions, and μ is a particular member of such a subclass – this will become clearer in the next section. We can now write two different definitions of the $\hat{H}_0$ operator, namely

$$\hat{H}_0^{\mathrm{SC}} = \sum_m^{\mathrm{CAS}} |\Psi_m^{(0)}\rangle E_m^{(0)} \langle \Psi_m^{(0)}| + \sum_{kl} |\Psi_l^{\{k\}}\rangle E_l^{\{k\}} \langle \Psi_l^{\{k\}}| \tag{10.135}$$

and

$$\hat{H}_0^{\mathrm{PC}} = \sum_m^{\mathrm{CAS}} |\Psi_m^{(0)}\rangle E_m^{(0)} \langle \Psi_m^{(0)}| + \sum_{kl\mu} |\Psi_{l\mu}^{\{k\}}\rangle E_{l\mu}^{\{k\}} \langle \Psi_{l\mu}^{\{k\}}| \ , \tag{10.136}$$

for the strongly and partially contracted versions of NEVPT2, respectively. The first term involves all possible states of the CASSCF wave function, while the second term includes the normalized perturbers of the first-order interacting space. Since, after the next section, we have these perturbers, we only need to find a way to obtain the associated "eigenenergies" $E_{l/l\mu}^{\{k\}}$. To do this, we will employ Dyall's Hamiltonian, $\hat{H}^D$ [35] (see below), in which the associated eigenvalues are simply computed as

$$E_l^{\{k\}} = \langle \Psi_l^{\{k\}} | \hat{H}^D | \Psi_l^{\{k\}} \rangle \tag{10.137}$$

for the strongly contracted case. For the partially contracted case a more elaborate procedure is engaged. First a projected Hamiltonian, based on $\hat{H}^D$, is formed with eigenfunctions and eigenvalues given by

$$\hat{P}_l^{\{k\}} \hat{H}^D \hat{P}_l^{\{k\}} \Psi_{l\mu}^{\{k\}} = E_{l\mu}^{\{k\}} \Psi_{l\mu}^{\{k\}}, \tag{10.138}$$

where $\Psi_{l\mu}^{\{k\}}$ are spanned by the linearly independent functions $\Phi_{lm}^{\{k\}}$ of the space $S_l^{\{k\}}$.

Before we continue, let us mention Dyall's Hamiltonian – here presented in a form invariant to unitary rotations inside each orbital class (inactive, active, and virtual). This Hamiltonian is introduced as an approximation to the electronic Hamiltonian and is subdivided, for computational efficiency, into an inactive–virtual part (iv) and an active part (act) as

$$\hat{H}^D = \hat{H}_{\mathrm{iv}} + \hat{H}_{\mathrm{act}} \ , \tag{10.139}$$

where the inactive–virtual part is a simple one-electron diagonal operator,

$$\hat{H}_{\mathrm{iv}} = \sum_{ij} f_{ij} \hat{E}_{ij} + \sum_{ab} f_{ab} \hat{E}_{ab} + C, \tag{10.140}$$

where we use the index definitions of Chapter 1, section 1.5. The constant C and the Fock matrix elements f_{pq} are defined below. For the active part of the operator, $\hat{H}_{\mathrm{act}}$, we have the following expression,

$$\hat{H}_{\mathrm{act}} = \sum_{tu} h_{tu}^{\mathrm{eff}} \hat{E}_{tu} + \frac{1}{2} \sum_{tuvx} (tu|vx) \hat{e}_{tuvx} \ , \tag{10.141}$$

where $h_{tu}^{\mathrm{eff}} = h_{tu} + \sum_i (2(tu|ii) - (ti|ui))$ are the effective one-electron integrals, introducing the field of the inactive electrons into the active part of the Hamiltonian. Furthermore, the constant C is chosen to be,

$$C = 2 \sum_i h_{ii} + \sum_{ij} [2(ii|jj) - (ij|ji)] - 2 \sum_k^{\mathrm{frozen}} \epsilon_k \ , \tag{10.142}$$

with which the result is equivalent to the full Hamiltonian within the CAS space,

$$\hat{H}^D \Psi_m^{(0)} = E_m^{(0)} \Psi_m^{(0)}. \tag{10.143}$$

Finally the Fock matrix elements f_{pq} are defined as

$$f_{ij} = -\langle \Psi_m^{(0)} | \hat{a}_i^\dagger \hat{H} \hat{a}_j | \Psi_m^{(0)} \rangle + \delta_{ij} E_m^{(0)} \tag{10.144}$$

and

$$f_{ab} = \langle \Psi_m^{(0)} | \hat{a}_a \hat{H} \hat{a}_b^\dagger | \Psi_m^{(0)} \rangle - \delta_{ab} E_m^{(0)}. \tag{10.145}$$

In this fashion the Dyall Hamiltonian is equivalent to the Møller–Plesset $\hat{H}_0$ in the inactive and the virtual space while in the active space it is the complete active space Hamiltonian. The latter ensures that the NEVPT2 approach is a size-consistent method.

10.3.5.2 The First-Order Interacting Space

Here we will describe the first-order interacting space (also denoted the SD-space) for the strongly contracted and the partially contracted schemes. In the SC-NEVPT2 version we will use only one perturber from each class in the first-order interacting space, while in the PC-NEVPT2 method we will include all possible states of the SD-space (just as in the CASPT2 method). The classes are similar to the classification in the CASPT2 method (see section 10.3.3.2 above). The index k, ranging from -2 to 2, represents the change in the number of electrons in the active space. An additional index, l, is used to discriminate between subclasses when there are several types of operators belonging to the same k-class. We note that we have eight different types of subclasses, just as in the CASPT2 method, although the primary classification scheme there was based on internal, semi-internal, and external excitations, instead of the k index.

In the SC-NEVPT2 method a single perturber per subclass is generated as follows,

$$\Psi_l^{\{k\}} = \hat{P}_l^{\{k\}} \hat{H} \Psi_m^{(0)}, \tag{10.146}$$

where $\Psi_m^{(0)}$ is the reference function of state m, and $\hat{P}_l^{\{k\}}$ is the projector onto $S_l^{\{k\}}$, the subspace of all states in subclass l of class k. To some extent this is similar to the first iteration in a CI solver algorithm. Since only one perturber is picked from each type of class there will be no problem with linear dependences.

For the PC-NEVPT2 method, however, we will employ all possible states in each subclass, just as in the CASPT2 method, and as in the CASPT2 case we will have possible linear dependences. Thus a similar procedure is adopted to generate a set of linearly independent functions, $\Phi_{lm}^{\{k\}}$, of subspace $S_l^{\{k\}}$. The final functions, $\Psi_{l\mu}^{\{k\}}$, to be used in the $\hat{H}_0$ operator, are generated by finding the eigenvalues and eigenvectors of the projection of Dyall's Hamiltonian on $S_l^{\{k\}}$ Eq. (10.138).

To sum up the technical description of the NEVPT2 formalism and highlight significant differences: (*i*) NEVPT2 uses a two-electron $\hat{H}_0$ resulting in longer computation time and the need of up to explicit four-particle density matrices, on the other hand, in NEVPT2 the Hamiltonian is actually block-diagonal for each excitation class, which reduces the size of the problem, (*ii*) the NEVPT2 method is strictly size extensive – CASPT2 is not, and (*iii*) the NEVPT2 method does not, from a practical perspective, suffer from intruder-state problems. The latter property is due to the fact that with the NEVPT2 partitioning of the Hamiltonian the energies of the perturbing wave functions – the states in the first-order interacting space – are far from degenerate with the energy of the zeroth-order state. The worst case is associated with an excited state corresponding to a semi-internal excitation which will effectively generate an energy denominator corresponding to the sum of the ionization potential of an active electron and the energy of a virtual orbital. The former should be negative and can at worst be close to zero for a Rydberg orbital, while the latter is positive (for neutral species) and rarely close to zero. Therefore, calculations for electronic

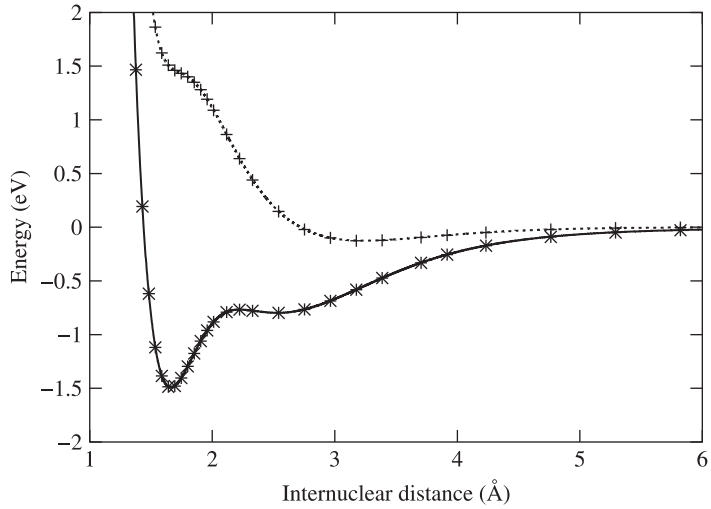


Figure 10.6 Potential energy curves for the $^1\Sigma^+$ states of Cr_2 (ANO basis set, (12/12) CAS-CI space). Dashed line (CASSCF) and full line (SC-NEVPT2). Reprinted with permission from ref. [19]. Copyright 2001, Elsevier.

excited states should be possible with NEVPT2 without experiencing any intruder state problem. For a detailed discussion on this we recommend the reader to consult ref. [41].

Finally, extensive benchmarking of the NEVPT2 method does not exist at present. However, let us again point out that different to CASPT2 the intruder state problem is effectively removed and the result is strictly size consistent. The former is explicitly demonstrated in fig. 10.6, in which the curve was generated without the need for any type of level-shifting technique. Preliminary benchmark studies by Angeli and co-workers [42] on valence and Rydberg transitions in formaldehyde and acetone indicated errors in the range of 0.1–0.2 eV and an absence of intruder states. A recent benchmark by Shapiro et al. [43] on vertical excitation energies of some 28 medium sized organic molecules, does not provide decisive arguments for any discrimination between the PC-NEVPT2, SC-NEVPT2, and SS-CASPT2 methods based on the accuracy of the predictions – if any, SS-CASPT2 has a small edge (see fig. 10.7). For example, the mean unsigned error is 0.28 eV, 0.23 eV, and 0.21 eV for the methods as ordered above. Rather than accuracy, it will be qualities such as computer efficiency and others that will, from case to case, be the determining factor in the selection of which method to use.

10.3.6 Performance Improvements

The use of the CASPT2 and NEVPT2 methods has been very successful in predicting and understanding a number of phenomena. However, the conventional implementations suffer from various bottlenecks such as the transformation of the integrals to the molecular orbital basis or the evaluation of the amplitudes. The former scales effectively as $\mathcal{O}(n^5)$ where n is the size of the basis set, while the latter suffers from the "exponential wall" typical of a conventional configuration-interaction treatment of the wave function. This limits the use of the methods to modest-sized systems with active spaces of up to about 18 electron in 18 orbitals.

For moderate active spaces, the integral transformation step tends to dominate. This integral-transformation bottleneck was effectively removed by the use of a CASPT2 implementation based on Cholesky decomposed (CD) integrals, the CD-CASPT2 approach [44]. This

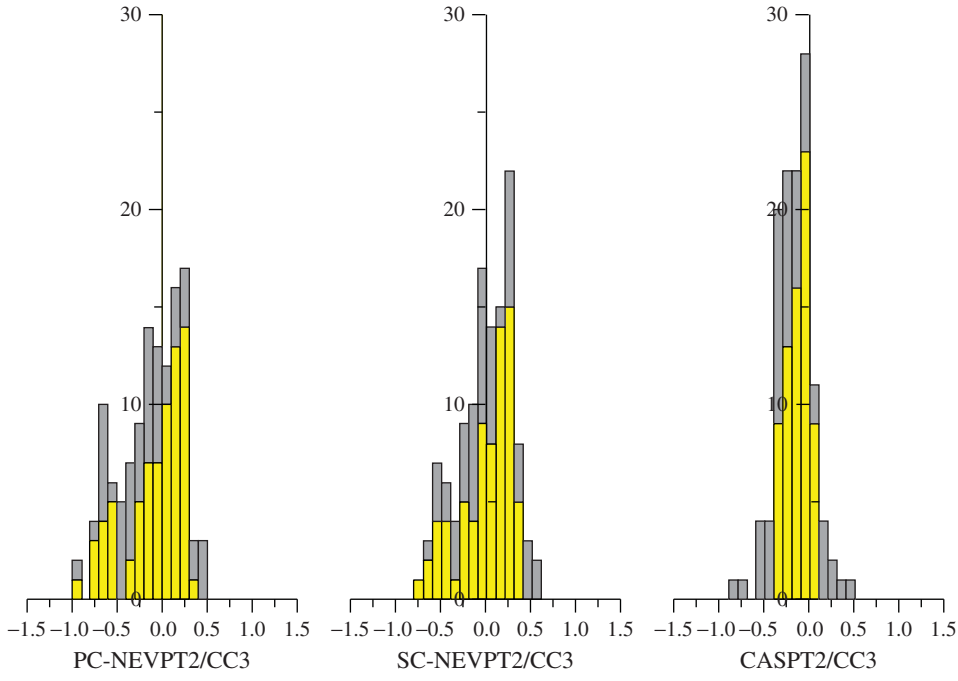


Figure 10.7 Histograms of the frequency of deviations in the singlet excitation energies in eV. Statistics of all excitation energies are shown in gray and those below 7 eV are represented by yellow bars. Reprinted with permission from ref. [43]. Copyright 2013, American Chemical Society.

reduced the scaling of the integral transformation by one order of magnitude while simultaneously reducing cache faults and I/O requests. Benchmarks demonstrate that CD in combination with CASSCF/CASPT2 reduces the computer time by a factor of four up to 1–2 orders of magnitude, with a loss of accuracy for excitation energies of organic molecules of about 0.001 eV [45]. Another early, but not as successful, approach is the frozen natural orbital CASPT2 (FNO-CASPT2) method [46, 47]. In this approach, an approximation to the one-particle density matrix corrected up to second order is diagonalized in the virtual–virtual block. Subsequently eigenvectors are eliminated if the corresponding eigenvalue is below a predefined threshold. It has been demonstrated that the use of this kind of reduced virtual space recovers some 95 % of the second-order correction to the energy while significantly reducing the computational expense. Some more recent and significant progress, however, has been made for both the CASPT2 and the NEVPT2 methods with respect to molecular-size scaling given a fixed active space – an ability which can be of significance in the study of transition metals with large ligands. In these approaches the effective number of amplitudes in the first-order interacting space is dramatically reduced without any significant loss of accuracy by a combined initial generation of projected atomic orbitals (PAO) [48] and a subsequent generation of pair natural orbital (PNO) [49]. Both the PNO-CASPT2 [50] and the domain-based local PNOs flavor of NEVPT2 (DBLPNO-NEVPT2) [51] methods have demonstrated, given a fixed active space, linear scaling with respect to molecular size in combination with insignificant loss of accuracy – recovering 99.9 % of the second-order correction to the energy. More recent developments using density matrix renormalization group (DMRG) approximations to treat large active spaces in CASPT2 [52–56] and NEVPT2 [57–59] – the DMRG-CASPT2 and DMRG-NEVPT2 methods – have been very successful in removing the

conventional limitation on the size of the active space. Finally, the progress in overcoming the bottleneck of the higher-order n-particle matrix evaluation should be mentioned [60–64]. In these studies, the efficiency of different approaches has been explored, with mixed results, namely: (*i*) approximating the cumulants of the higher-order reduced density matrices [60, 65], (*ii*) using a matrix product state representation (found, for example, in DMRG) in the computation of the first-order interacting space – called MPS-PT2 – in state-specific or quasi-degenerate (see below) PT [61, 62, 66], (*iii*) using a time-dependent formulation of MRPT [63], (*iv*) and a projection of the first-order interacting space in an SC-NEVPT2 setting onto a renormalized Hilbert space [64].

10.4 Quasi-Degenerate Perturbation Theory

In the course of using the SS-CASPT2 method, it was evident that the approach is only usable when the reference functions are close to qualitatively correct, and that the first-order correction to the wave function would be sufficient for a quantitatively correct energy only if the perturbation expansion terminates at second order. Anything else would be wishful thinking. In the literature there are ample cases in which the reference CASSCF wave functions are qualitatively wrong due to the lack of dynamic electron correlation – see, for example, the case of the V-state of ethylene, in which CASSCF introduces an artificial mixing of Rydberg and valence character, something the SS-CASPT2 approach never recovers from [67]. The trivial solution to this problem could have been to increase the active space or to go beyond second order in the perturbation series – which sounds trivial but is expensive, complicated, or not even feasible. A more pragmatic approach involves formulating a multi-state version of perturbation theory, something that is not straightforwardly done from the single-state theory. What is the problem? Well, the single-state approach has the problem that, if applied to several states at the same time, it does not allow the reference states to trivially mix with each other as the perturbation is introduced order by order. To avoid this, multi-state perturbation theories use an effective Hamiltonian procedure to produce corrected energies originating from corrected *model states*, which are linear combinations of the original *reference states*. This gives rise to quasi-degenerate perturbation theory (QDPT), the basics of which we will see in this section.

To proceed with a multi-state theory, however, some new tools have to be derived – the Bloch equations and the wave operator (see, for example, refs. [68–70]). We start by defining a model space – or reference space – commonly denoted P, spanned by some eigenfunctions of the Hamiltonian of the unperturbed system $\hat{H}_0$, the *reference states*,

$$\Psi_i^{(0)} \quad i = 1, \dots, d, \tag{10.147}$$

where d is the dimension of the model space. We also define the operators

$$\hat{P} = \sum_{i=1}^{d} |\Psi_i^{(0)}\rangle\langle\Psi_i^{(0)}| \text{ , and} \tag{10.148}$$

$$\hat{Q} = 1 - \hat{P}. \tag{10.149}$$

The operator $\hat{P}$ can now be used to project the exact wave function of state i onto the the model space as

$$\hat{P}\Psi_i = \Psi_i^P = \sum_{j=1}^{d} c_{ij}\Psi_j^{(0)}, \tag{10.150}$$

where Ψ_i^P is the *model state* of state Ψ_i, expressed by a linear combination of the reference states in the model space. These linear combinations are found by diagonalization of the multi-state Hamiltonian matrix. The wave operator, $\hat{\Omega}$, does the corresponding back-transformation, acting on the model states it generates the exact states,

$$\Psi_i = \hat{\Omega}\Psi_i^P \quad i = 1, \dots, d, \tag{10.151}$$

where we note the similarities to Eq. (10.19) for the single-state case – both provide a recipe to find the corrections of the wave function(s) to arbitrary order. For completeness, $\hat{\Omega}$ is defined as a null operator when operating on the complementary space, i.e., $\hat{\Omega}\hat{Q} = 0$ and therefore $\hat{\Omega}\hat{P} = \hat{\Omega}(1 - \hat{Q}) = \hat{\Omega}$.

We will here derive the Bloch equation – the master equation to express the wave operator – in which an explicit dependence on the exact energy is avoided. First, we start by left-multiplying the "exact" Schrödinger equation with $\hat{\Omega}\hat{P}$, that is,

$$\hat{\Omega}\hat{P}(\hat{H}_0 + V)\Psi_i = E_i\Psi_i \ . \tag{10.152}$$

The operator $\hat{\Omega}\hat{P}$ will project any exact wave function to the reference space and then subsequently regenerate the exact wave function. Hence, the operator will leave an exact wave function untouched, and the RHS of the equation is identical to the original expression. Using $[\hat{H}_0, \hat{P}] = 0$, and with the help of Eqs. (10.150) and (10.151), the following expression is found

$$(\hat{\Omega}\hat{H}_0 + \hat{\Omega}\hat{P}V\hat{\Omega})\Psi_i^P = E_i\Psi_i \ , \tag{10.153}$$

and by using Eq. (10.151), $\hat{\Omega}\hat{P} = \hat{\Omega}$, and the original Schrödinger equation to eliminate the energy, one gets

$$(\hat{\Omega}\hat{H}_0 + \hat{\Omega}V\hat{\Omega})\Psi_i^P = (\hat{H}_0\hat{\Omega} + V\hat{\Omega})\Psi_i^P, \tag{10.154}$$

which holds for the model space. Since the operators on both sides are null when operating on the complementary space, it holds for any state, and can be paraphrased into the generalized form of the Bloch equation,

$$[\hat{\Omega}, \hat{H}_0] = V\hat{\Omega} - \hat{\Omega}V\hat{\Omega}. \tag{10.155}$$

In the single-state case the equation can be rewritten as

$$(E_i^{(0)} - \hat{H}_0)\hat{\Omega} = V\hat{\Omega} - \hat{\Omega}V\hat{\Omega}. \tag{10.156}$$

Expanding the wave operator in powers of the perturbation with $\hat{\Omega}^{(0)} = \hat{P}$ [68] we have

$$\hat{\Omega} = \hat{P} + \hat{\Omega}^{(1)} + \hat{\Omega}^{(2)} + \dots, \tag{10.157}$$

and finally, due to the recursive nature of the Bloch equation we have

$$[\hat{\Omega}^{(n)}, \hat{H}_0] = \hat{Q}V\hat{\Omega}^{(n-1)} - \sum_{k=1}^{n-1} \hat{\Omega}^{(n-k)}V\hat{\Omega}^{(k-1)} \ . \tag{10.158}$$

Then we get the wave operator, up to third order, as

$$[\hat{\Omega}^{(1)}, \hat{H}_0] = \hat{Q}V\hat{P} = \hat{Q}\hat{H}\hat{P}, \tag{10.159}$$

$$[\hat{\Omega}^{(2)}, \hat{H}_0] = \hat{Q}V\hat{\Omega}^{(1)} - \hat{\Omega}^{(1)}V\hat{P}, \text{ and} \tag{10.160}$$

$$[\hat{\Omega}^{(3)}, \hat{H}_0] = \hat{Q}V\hat{\Omega}^{(2)} - \hat{\Omega}^{(2)}V\hat{P} - \hat{\Omega}^{(1)}V\hat{\Omega}^{(1)} \ . \tag{10.161}$$

In the single-state case the LHS is replaced with $(E_i^0 - \hat{H}_0)\hat{\Omega}_i^{(n)}$, where the wave operator is unique to the state i, and the general wave operator is now expressed as

$$\hat{\Omega}^{(n)} = \sum_{i=1}^{d} \hat{\Omega}_i^{(n)} |\Psi_i^{(0)}\rangle\langle\Psi_i^{(0)}| \,. \tag{10.162}$$

The effective Hamiltonian, by definition, is written as

$$\hat{H}^{\text{eff}} = \hat{P}\hat{H}\hat{\Omega}\hat{P}, \tag{10.163}$$

which we note only acts on the reference space. The model states, $\Psi_i^{(0)}$, $i = 1, \ldots, d$, are eigenfunctions of the effective Hamiltonian,

$$\hat{H}^{\text{eff}}\Psi_i^{(0)} = E_i\Psi_i^{(0)}, \tag{10.164}$$

and the effective Hamiltonian up to second order is

$$\hat{H}^{\text{eff[2]}} = \hat{P}\hat{H}\hat{P} + \hat{P}\hat{H}\hat{\Omega}^{(1)}\hat{P}. \tag{10.165}$$

We note at once that the energy corrected up to first order is computed as

$$E_i^{[1]} = \langle\Psi_i^{(0)}|\hat{H}_0 + V|\Psi_i^{(0)}\rangle \,, \tag{10.166}$$

and the second-order correction to the energy is expressed as

$$E_i^{(2)} = \langle\Psi_i^{(0)}|\hat{H}|\Psi_i^{(1)}\rangle \,. \tag{10.167}$$

Let us look at the first-order corrected wave function in terms of a single- and multi-state formalism. The first-order correction to the wave function in a single-state formalism is expressed as

$$\Psi_i^{(1)} = \hat{\Omega}_i^{(1)}\Psi_i^{(0)} = \sum_{\substack{j=1 \\ j\neq i}}^{\infty} C_{ij}\Psi_j^{(0)}, \tag{10.168}$$

while in a multi-state approach we have (see Eqs. 10.150 and 10.162)

$$\begin{aligned} \Psi_i^{(1),P} &= \hat{\Omega}^{(1)}\Psi_i^P \qquad (10.169)\\ &= \sum_{j=1}^{d} \hat{\Omega}_j^{(1)} |\Psi_j^{(0)}\rangle\langle\Psi_j^{(0)}|\Psi_i^P\rangle \\ &= \sum_{j=1}^{d} c_{ji} \sum_{\substack{k=1 \\ k\neq j}}^{\infty} C_{jk}\Psi_k^{(0)} \\ &= \sum_{j=1}^{d} c_{ji}\Psi_j^{(1)}. \qquad (10.170) \end{aligned}$$

Note that this does not imply that the first-order corrected model states are orthogonal, they will only be so at infinite order. The coefficients, C_{jk}, are computed from the general Bloch equation according to

$$\sum_{j\neq i} C_{ij}\langle\Psi_k^{(0)}|(E_i^{(0)} - \hat{H}_0)|\Psi_j^{(0)}\rangle = \langle\Psi_k^{(0)}|\hat{H}|\Psi_i^{(0)}\rangle \,, \tag{10.171}$$

which simplifies to

$$C_{ik} = \frac{\langle\Psi_k^{(0)}|V|\Psi_i^{(0)}\rangle}{E_i^{(0)} - E_k^{(0)}} \,, \tag{10.172}$$

and now allows the computation of the second-order energy as described above. The matrix elements of the effective Hamiltonian are now expressed as

$$\langle \Psi_i^{(0)} | \hat{H}^{\text{eff}[2]} | \Psi_i^{(0)} \rangle = E_i^{[2]} \tag{10.173}$$

and

$$\langle \Psi_i^{(0)} | \hat{H}^{\text{eff}[2]} | \Psi_j^{(0)} \rangle = \langle \Psi_i^{(0)} | \hat{H} | \Psi_j^{(0)} + \Psi_j^{(1)} \rangle \, . \tag{10.174}$$

Given these matrix elements we can now proceed with a diagonalization and derive the model states of the model space. To conclude, this approach can be viewed as *perturb-then-diagonalize*, in which we use perturbation theory to set up the effective Hamiltonian which is subsequently diagonalized.

10.5 Multi-State Multi-Configurational Reference Perturbation Methods

In the pursuit of practical methods to overcome single-state MRPT limitations, the tools of quasi-degenerate perturbation theory come in handy. Defining a relevant model space of several states provides a pathway in which (*i*) the computational effort only increases linearly with the number of states in the model space and (*ii*) artificial redundancies (due to deficiencies on the CASSCF treatment) are removed, in one single step, under the influence of the perturbation (that is, including electron–electron correlation). These methods will now be of the nature *diagonalize-then-perturb-then-diagonalize* approach, in the sense that we will use basis functions in the setup of the effective Hamiltonian which are based on a *diagonalize-then-perturb* technique. Specifically, below we will describe two such methods in some detail – the multi-state and the extended multi-state CASPT2 (MS-CASPT2 and XMS-CASPT2). Similar approaches have successfully been applied to both the MRMP2 (multi-configurational quasi-degenerate second-order perturbation theory, MCQDPT2) [71, 72] and NEVPT2 single-state theories (QD-NEVPT2) [73].

10.5.1 Multi-State CASPT2 Theory

As described above, QDPT requires that $\hat{H}_0$ and $\hat{H}_i^{\text{eff}[2]}$ be unique for all the reference states. However, much of the success of SS-CASPT2 lies in the use of an $\hat{H}_0$ that is specifically tailored to the state under study. In 1998 Finley and co-workers suggested a hybrid approach of SS-CASPT2 with quasi-degenerate perturbation theory [74] – the multi-state CASPT2 (MS-CASPT2) method – to improve the SS-CASPT2 method in those cases where the reference CASSCF wave function does not meet the qualitative requirements for a fast convergence of the perturbation expansion, additionally it should remove erroneous behavior of SS-CASPT2 at points of degeneracy. This approach is very similar to what was described above for standard quasi-degenerate perturbation theory, that is a *perturb-then-diagonalize* approach. However, in this case the starting reference functions have already been the subject of a diagonalization procedure – the CASSCF procedure – and in this aspect the MS-CASPT2 procedure is a *diagonalize-then-perturb-then-diagonalize*-like procedure. The suggestion of Finley et al. revolves entirely around Eqs. (10.173) and (10.174). How are those equations to be used in association with the SS-CASPT2 approach? In this respect the authors suggest the following. First, and maybe trivial, the diagonal elements in the effective Hamiltonian are simply the individual SS-CASPT2 energies,

$$\langle \Psi_i^{(0)} | \hat{H}^{\text{MS-CASPT2}} | \Psi_i^{(0)} \rangle = E_i^{\text{SS-CASPT2}} = \langle \Psi_i^{(0)} | \hat{H}_i^{\text{eff}[2]} | \Psi_i^{(0)} \rangle \, . \tag{10.175}$$

At this point we also recognize the significant difference to Eq. (10.173): the effective Hamiltonian operator and the corresponding implicit partitioning of the Hamiltonian is *state-specific* [75]. This raises a problem when we proceed to the computation of the off-diagonal elements of the effective Hamiltonian matrix, that is,

$$\langle \Psi_i^{(0)} | \hat{H}_j^{\text{eff}[2]} | \Psi_j^{(0)} \rangle = \langle \Psi_i^{(0)} | \hat{H} | \Psi_j^{(1)} \rangle \neq \langle \Psi_j^{(0)} | \hat{H} | \Psi_i^{(1)} \rangle = \langle \Psi_j^{(0)} | \hat{H}_i^{\text{eff}[2]} | \Psi_i^{(0)} \rangle \ , \tag{10.176}$$

which for obvious reason produces a non-symmetric Hamiltonian matrix. In the current implementation this problem is fixed with a simple symmetrization of the matrix with the limitation that results should be questioned if corresponding off-diagonal elements are large and different. The MS-CASPT2 results are generally in line with SS-CASPT2 results. In the original report by Finley and co-workers, the authors demonstrated explicitly how the MS-CASPT2 approach solves the problem with the artificial valence–Rydberg mixing in the V-state of ethane, and how MS-CASPT2, unlike SS-CASPT2, provides qualitatively correct results around the avoided crossing of the ionic–neutral states in the dissociation of LiF (see fig. 10.8). When MS-CASPT2 and SS-CASPT2 results differ, however, the prudent user will have to address the issue of whether or not the MS-CASPT2 results should be trusted. In particular, it has been demonstrated that the MS-CASPT2 method resolves cases with accidental degeneracy at the CASSCF level if it does not persist at the SS-CASPT2 level. On the other hand, the MS-CASPT2 is known to provide erroneous behavior around conical intersections [76, 77] and other cases of near-degeneracy. The latter problem is known to be a consequence of the use of state-specific effective operators, a procedure will be suggested below to remove these problems.

10.5.2 Extended MS-CASPT2 Theory

Granovsky analysed MCQDPT theory and put forward qualitative criteria which a quasi-degenerate perturbation theory should satisfy [77]. One of these is that the eigenvalues of the effective Hamiltonian should be invariant to the particular basis functions selected to span the model space, as long as they span the same subspace. In this respect Granovsky suggested that the $\hat{H}_0$ operator should be extended from

$$\hat{H}_0 = \sum_i |\Psi_i^{(0)}\rangle\langle\Psi_i^{(0)}|\hat{F}|\Psi_i^{(0)}\rangle\langle\Psi_i^{(0)}| \tag{10.177}$$

to

$$\hat{H}_0 = \sum_{ij} |\Psi_i^{(0)}\rangle\langle\Psi_i^{(0)}|\hat{F}|\Psi_j^{(0)}\rangle\langle\Psi_j^{(0)}| \ , \tag{10.178}$$

where $\hat{F}$ is a Fock-like operator, as in CASPT2. Thus, the $\hat{H}_0$ operator is explicitly independent of the selection of the basis functions spanning a particular model space. This new technique is denoted as extended MCQDPT (XMCQDPT). This extension has been applied to MS-CASPT2 theory yielding the method known as XMS-CASPT2 [78]. In this approach the $\hat{H}_0$ matrix is diagonalized. The "non-extended" procedure is then continued in the basis of these eigenfunctions $\Psi_i^{X(0)}$.

It should be noted that the XMS-CASPT2 approach requires the use of a state-specific partitioning to be abandoned, as used in the MS-CASPT2 method. The state-specific approach and the MS-CASPT2 method have been successful in predicting vertical excitation energies. The

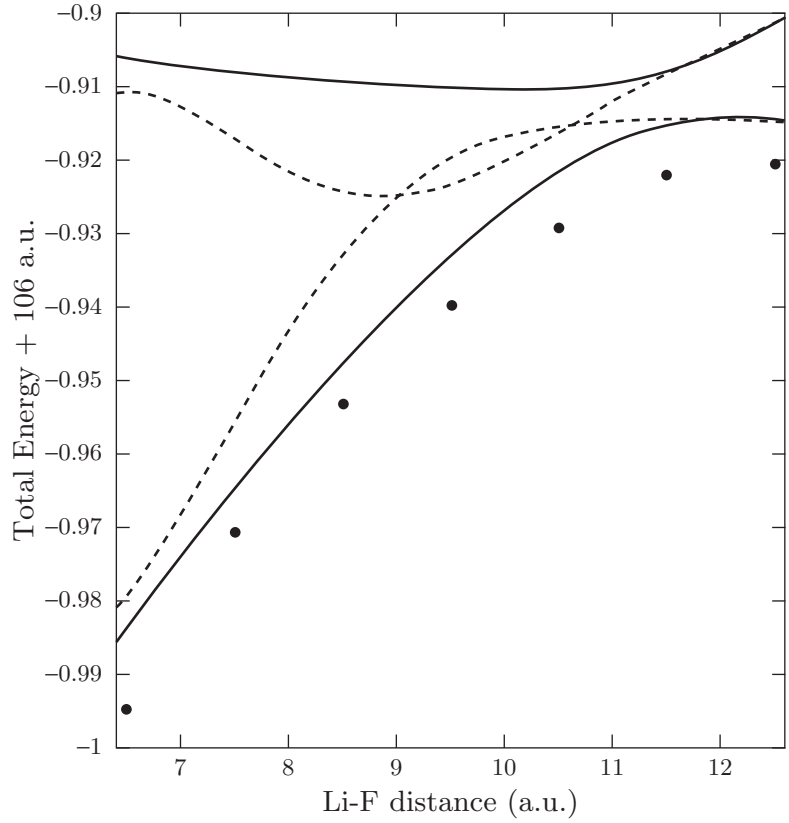


Figure 10.8 Potential energy curves for the two lowest $^1\Sigma^+$ states of LiF in the region of their neutral–ionic avoided crossing. The dotted lines represent the SS-CASPT2 calculations and the solid lines represent the MS-CASPT2 calculations. The points indicate the FCI calculations. Reprinted with permission from ref. [74]. Copyright 1998, Elsevier.

XMS-CASPT2 method certainly will resolve some of the flaws of the MS-CASPT2 theory in and close to conical intersections (see fig. 10.9). However, this will be at the expense of accuracy of predicting vertical excitation energies for well-separated states. Furthermore, in the XMS-CASPT2 method the eigenvalues of the model space are sometimes too sensitive to changes of the model space by including more or fewer states.

10.6 Summary and Outlook

In this chapter we have described the CASPT2, MS-CASPT2, and XMS-CASPT2 methods in some detail. We have also lightly touched on some related methods such as the CAS-MP2, MRMP2, and NEVPT2 approximations. This has been presented in a bottom-up order, starting with a thorough presentation of Rayleigh–Schrödinger perturbation theory. Here we also discussed the issue of the

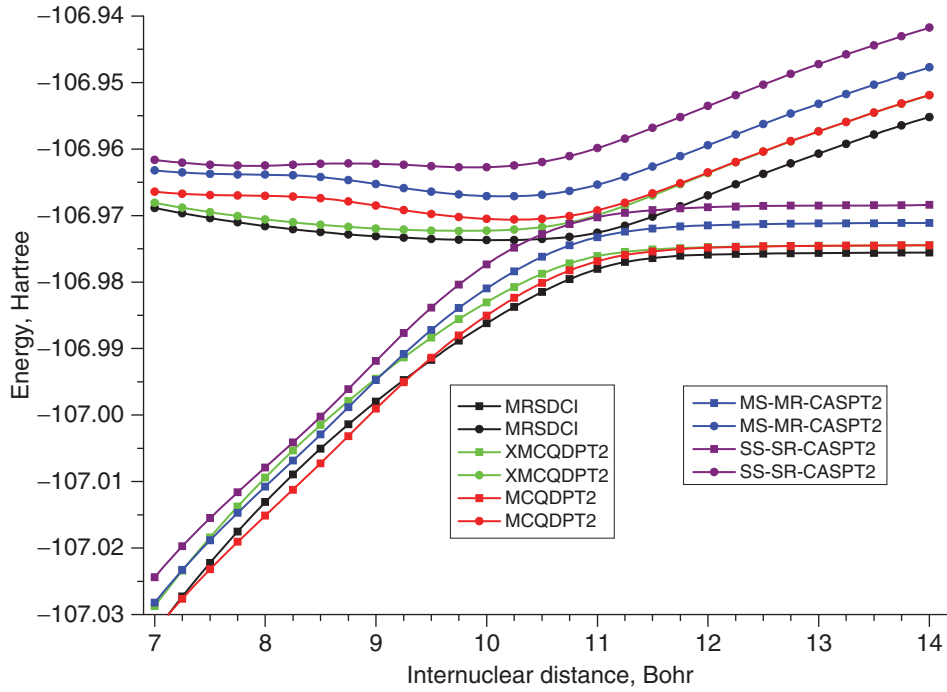


Figure 10.9 Potential energy curves (PECs) of the LiF molecule near the avoided crossing area computed using different MS-MRPT approaches. MRSDCI PECs are given as reference. PECs were intentionally left non-shifted to a common dissociation limit. Reprinted with permission from ref. [77]. Copyright 2011, American Institute of Physics.

convergence of the perturbation expansion and how divergence or near-divergence is manifested as the presence of intruder states. In the analysis of these we introduced the gap-shift technique. This tool later became a general technique to avoid divergence – real and imaginary shift techniques – or a way to compensate for systematic errors – the IPEA shift. This was followed up by a detailed presentation of the Møller–Plesset partitioning in association with the single determinant SCF wave function – the CASPT2 model is supposed to reduce to this as the active space is reduced to null. Here we also discussed numerical methods to compute the first-order correction to the wave function and subsequently the second-order correction to the energy, when analytical solutions do not exist. At this point the single-state multi-configurational perturbation schemes were introduced (CAS-MP2, CASPT2, MRMP2, and NEVPT2). The chapter discussed the benchmark results of these methods, which clearly exhibit explicit or implicit properties of the single-state approaches. These second-order perturbation approaches have their limitations and multi-state versions of perturbation theory have, sometimes, to be used to relieve these shortcomings. For that purpose we presented, in some detail, the basic steps of quasi-degenerate perturbation theory – this allowed us to pave the way for an understanding of the multi-state versions of perturbation theories mentioned above. Prepared by this introduction, we proceeded with a description of the multi-state version of CASPT2, which has some parallels with the corresponding MRMP2 and NEVPT2 versions. Finally, an extended version of MS-CASPT2 was introduced – based on extended quasi-degenerate perturbation theory – which will have an $\hat{H}_0$ operator whose eigenvalues are invariant to rotations of the

states in the model space. This, in particular, should resolve some problems of MS-CASPT2 with artificial surface crossings observed close to or at conical intersections.

The future of MRPT, its implementation and approximations, is closely connected with the partitioning of the Hamiltonian for the CASPT2 method. Here there is still room for significant improvement. While the NEVPT2 model does not suffer from the intruder-state problem to the degree that CASPT2 does, it derives this character at a larger computational cost. For the NEVPT2 approach to become tractable for larger active spaces, approximations will have to be developed to reduce this disadvantage compared to the CASPT2 method. Additionally, with the advent of the DMRG approach to enable CASSCF calculations with large active spaces, both the CASPT2 and the NEVPT2 methods have to be recast to overcome the issues that computing and storing the whole (or partially contracted) four-particle density matrix – at least if the notion of contracted reference functions is to survive. For recent developments in MRPT, we point the reader to the references mentioned at the end of section 10.3.6. Finally, the issue of how to select the reference functions needs to be readdressed. Is there maybe a golden path to be explored in which the reference function already – through some simple approach – carries a significant part of the dynamic electron correlation?

Acknowledgments

The authors acknowledge fruitful discussions with Kerstin Andersson, Celestino Angeli, Leon Freitag, Kimihiko Hirao, and Markus Reiher, and funding from the Swedish Research Council (grant 2016-03398).

References

1 Schrödinger, E. (1926). Quantisierung als Eigenwertproblem. *Ann. Phys.* 385 (13): 437–490.
2 Strutt, B.R. and William, J. (1894). *The Theory of Sound*, 2e. London: MacMillan.
3 Helgaker, T., Jørgensen, P., and Olsen, J. (2000). *Molecular Electronic-Structure Theory*. Wiley.
4 Kato, T. (1995). *Perturbation Theory for Linear Operators*. Berlin Heidelberg: Springer.
5 Olsen, J., Jørgensen, P., Helgaker, T., and Christiansen, O. (2000). Divergence in Møller–Plesset theory: a simple explanation based on a two-state model. *J. Chem. Phys.* 112 (22): 9736–9748.
6 Roos, B.O. and Andersson, K. (1995). Multiconfigurational perturbation theory with level shift – the Cr_2 potential revisited. *Chem. Phys. Lett.* 245 (2–3): 215–223.
7 Forsberg, N. and Malmqvist, P.-Å. (1997). Multiconfiguration perturbation theory with imaginary level shift. *Chem. Phys. Lett.* 274 (1–3): 196–204.
8 Surján, P.R. and Szabados, Á. (1996). Damping of perturbation corrections in quasidegenerate situations. *J. Chem. Phys.* 104 (9): 3320–3324.
9 Epstein, P.S. (1926). The Stark effect from the point of view of Schroedinger's quantum theory. *Phys. Rev.* 28 (4): 695–710.
10 Nesbet, R.K. (1955). Configuration interaction in orbital theories. *Proc. Math. Phys. Eng. Sci.* 230 (1182): 312–321.

11 McDouall, J.J.W., Peasley, K., and Robb, M.A. (1988). A simple MC SCF perturbation theory: orthogonal valence bond Møller–Plesset 2 (OVB MP2). *Chem. Phys. Lett.* 148 (2–3): 183–189.

12 Andersson, K., Malmqvist, P.Å., Roos, B.O. et al. (1990). Second-order perturbation theory with a CASSCF reference function. *J. Phys. Chem.* 94 (14): 5483–5488.

13 Andersson, K., Malmqvist, P.-Å., and Roos, B.O. (1992). Second-order perturbation theory with a complete active space self-consistent field reference function. *J. Chem. Phys.* 96 (2): 1218–1226.

14 Hirao, K. (1992). Multireference Møller–Plesset method. *Chem. Phys. Lett.* 190 (3–4): 374–380.

15 Hirao, K. (1992). Multireference Møller–Plesset perturbation theory for high-spin open-shell systems. *Chem. Phys. Lett.* 196 (5): 397–403.

16 Hirao, K. (1992). Multireference Møller–Plesset perturbation treatment of potential energy curve of N_2. *Int. J. Quantum Chem.* 44 (S26): 517–526.

17 Hirao, K. (1993). State-specific multireference Møller–Plesset perturbation treatment for singlet and triplet excited states, ionized states and electron attached states of H_2O. *Chem. Phys. Lett.* 201 (1–4): 59–66.

18 Angeli, C., Cimiraglia, R., Evangelisti, S. et al. (2001). Introduction of *n*-electron valence states for multireference perturbation theory. *J. Chem. Phys.* 114 (23): 10252–10264.

19 Angeli, C., Cimiraglia, R., and Malrieu, J.-P. (2001). *n*-electron valence state perturbation theory: a fast implementation of the strongly contracted variant. *Chem. Phys. Lett.* 350 (3–4): 297–305.

20 Angeli, C., Cimiraglia, R., and Malrieu, J.-P. (2002). *n*-electron valence state perturbation theory: a spinless formulation and an efficient implementation of the strongly contracted and of the partially contracted variants. *J. Chem. Phys.* 117 (20): 9138–9153.

21 Lischka, H., Nachtigallová, D., Aquino, A.J.A. et al. (2018). Multireference approaches for excited states of molecules. *Chem. Rev.* 118 (15): 7293–7361.

22 Löwdin, P.-O. (1951). A note on the quantum-mechanical perturbation theory. *J. Chem. Phys.* 19 (11): 1396–1401.

23 Andersson, K. and Roos, B.O. (1995). Multiconfigurational second-order perturbation theory. In: *Modern Electronic Structure Theory*, 55–109. World Scientific Publishing Company.

24 Andersson, K. (1995). Different forms of the zeroth-order Hamiltonian in second-order perturbation theory with a complete active space self-consistent field reference function. *Theor. Chim. Acta* 91 (1–2): 31–46.

25 Malmqvist, P.Å., Pierloot, K., Shahi, A.R.M. et al. (2008). The restricted active space followed by second-order perturbation theory method: theory and application to the study of CuO_2 and Cu_2O_2 systems. *J. Chem. Phys.* 128 (20): 204109.

26 Ma, D., Manni, G.L., Olsen, J., and Gagliardi, L. (2016). Second-order perturbation theory for generalized active space self-consistent-field wave functions. *J. Chem. Theory Comput.* 12 (7): 3208–3213.

27 Andersson, K. and Roos, B.O. (1993). Multiconfigurational second-order perturbation theory: a test of geometries and binding energies. *Int. J. Quantum Chem.* 45 (6): 591–607.

28 Guner, V., Khuong, K.S., Leach, A.G. et al. (2003). A standard set of pericyclic reactions of hydrocarbons for the benchmarking of computational methods: the performance of ab initio, density functional, CASSCF, CASPT2, and CBS-QB3 methods for the prediction of activation barriers, reaction energetics, and transition state geometries. *Chem. A Eur. J.* 107 (51): 11445–11459.

29 Schreiber, M., Silva-Junior, M.R., Sauer, S.P.A., and Thiel, W. (2008). Benchmarks for electronically excited states: CASPT2, CC2, CCSD, and CC3. *J. Chem. Phys.* 128 (13): 134110.

30 Serrano-Andrés, L. (2011). Organic spectroscopy under Björn O. Roos. *Int. J. Quantum Chem.* 111 (13): 3284–3290.

31 Pierloot, K. (2011). Transition metals compounds: outstanding challenges for multiconfigurational methods. *Int. J. Quantum Chem.* 111 (13): 3291–3301.

32 Gagliardi, L. (2011). The study of actinide chemistry with multiconfigurational quantum chemical methods. *Int. J. Quantum Chem.* 111 (13): 3302–3306.

33 Budzák, i., Scalmani, G., and Jacquemin, D. (2017). Accurate excited-state geometries: a CASPT2 and coupled-cluster reference database for small molecules. *J. Chem. Theory Comput.* 13 (12): 6237–6252.

34 Ghigo, G., Roos, B.O., and Malmqvist, P.-Å. (2004). A modified definition of the zeroth-order Hamiltonian in multiconfigurational perturbation theory (CASPT2). *Chem. Phys. Lett.* 396 (1–3): 142–149.

35 Dyall, K.G. (1995). The choice of a zeroth-order Hamiltonian for second-order perturbation theory with a complete active space self-consistent-field reference function. *J. Chem. Phys.* 102 (12): 4909–4918.

36 Aquilante, F., Autschbach, J., Carlson, R.K. et al. (2015). Molcas8: new capabilities for multiconfigurational quantum chemical calculations across the periodic table. *J. Comput. Chem.* 37 (5): 506–541.

37 Kepenekian, M., Robert, V., and Le Guennic, B. (2009). What zeroth-order Hamiltonian for CASPT2 adiabatic energetics of Fe(II)N_6 architectures? *J. Chem. Phys.* 131 (11): 114702.

38 Zobel, J.P., Nogueira, J.J., and González, L. (2017). The IPEA dilemma in CASPT2. *Chem. Sci.* 8 (2): 1482–1499.

39 Vela, S., Fumanal, M., Ribas-Ariño, J., and Robert, V. (2015). On the zeroth-order Hamiltonian for CASPT2 calculations of spin crossover compounds. *J. Comput. Chem.* 37 (10): 947–953.

40 Daku Lawson, L.M., Aquilante, F., Robinson, T.W., and Hauser, A. (2012). Accurate spin-state energetics of transition metal complexes. 1. CCSD(T), CASPT2, and DFT study of $[M(NCH)_6]^{2+}$ (M = Fe, Co). *J. Chem. Theory Comput.* 8 (11): 4216–4231.

41 Angeli, C. and Cimiraglia, R. (2014). Some useful odds and ends from the *n*-electron valence state perturbation theory. *Chem. A Eur. J.* 118 (33): 6435–6439.

42 Angeli, C., Borini, S., and Cimiraglia, R. (2004). An application of second-order *n*-electron valence state perturbation theory to the calculation of excited states. *Theor. Chem. Acc.* 111 (2–6): 352–357.

43 Schapiro, I., Sivalingam, K., and Neese, F. (2013). Assessment of *n*-electron valence state perturbation theory for vertical excitation energies. *J. Chem. Theory Comput.* 9 (8): 3567–3580.

44 Aquilante, F., Malmqvist, P.-Å., Pedersen, T.B. et al. (2008). Cholesky decomposition-based multiconfiguration second-order perturbation theory (CD-CASPT2): application to the spin-state energetics of Co^{III}(diiminato)(NPh). *J. Chem. Theory Comput.* 4 (5): 694–702.

45 Boström, J., Delcey, M.G., Aquilante, F. et al. (2010). Calibration of Cholesky auxiliary basis sets for multiconfigurational perturbation theory calculations of excitation energies. *J. Chem. Theory Comput.* 6 (3): 747–754.

46 Aquilante, F., Todorova, T.K., Gagliardi, L. et al. (2009). Systematic truncation of the virtual space in multiconfigurational perturbation theory. *J. Chem. Phys.* 131 (3): 034113.

47 Segarra-Martí, J., Garavelli, M., and Aquilante, F. (2015). Multiconfigurational second-order perturbation theory with frozen natural orbitals extended to the treatment of photochemical problems. *J. Chem. Theory Comput.* 11 (8): 3772–3784.

48 Pulay, P. (1983). Localizability of dynamic electron correlation. *Chem. Phys. Lett.* 100 (2): 151–154.

49 Fink, R. and Staemmler, V. (1993). A multi-configuration reference CEPA method based on pair natural orbitals. *Theor. Chim. Acta* 87 (1–2): 129–145.

50 Menezes, F., Kats, D., and Werner, H.-J. (2016). Local complete active space second-order perturbation theory using pair natural orbitals (PNO-CASPT2). *J. Chem. Phys.* 145 (12): 124115.

51 Guo, Y., Sivalingam, K., Valeev, E.F., and Neese, F. (2016). SparseMaps–A systematic infrastructure for reduced-scaling electronic structure methods. III. Linear-scaling multireference domain-based pair natural orbital n-electron valence perturbation theory. *J. Chem. Phys.* 144 (9): 094111.

52 Kurashige, Y. and Yanai, T. (2011). Second-order perturbation theory with a density matrix renormalization group self-consistent field reference function: theory and application to the study of chromium dimer. *J. Chem. Phys.* 135 (9): 094104.

53 Liu, F., Kurashige, Y., Yanai, T., and Morokuma, K. (2013). Multireference ab initio density matrix renormalization group (DMRG)-CASSCF and DMRG-CASPT2 study on the photochromic ring opening of spiropyran. *J. Chem. Theory Comput.* 9 (10): 4462–4469.

54 Kurashige, Y. (2013). Multireference electron correlation methods with density matrix renormalisation group reference functions. *Mol. Phys.* 112 (11): 1485–1494.

55 Wouters, S., Van Speybroeck, V., and Van Neck, D. (2016). DMRG-CASPT2 study of the longitudinal static second hyperpolarizability of all-trans polyenes. *J. Chem. Phys.* 145 (5): 054120.

56 Nakatani, N. and Guo, S. (2017). Density matrix renormalization group (DMRG) method as a common tool for large active-space CASSCF/CASPT2 calculations. *J. Chem. Phys.* 146 (9): 094102.

57 Guo, S., Watson, M.A., Hu, W. et al. (2016). *n*-electron valence state perturbation theory based on a density matrix renormalization group reference function, with applications to the chromium dimer and a trimer model of poly(*p*-phenylenevinylene). *J. Chem. Theory Comput.* 12 (4): 1583–1591.

58 Knecht, S., Hedegård, E.D., Keller, S. et al. (2016). New approaches for *ab initio* calculations of molecules with strong electron correlation. *CHIMIA Int. J. Chem.* 70 (4): 244–251.

59 Freitag, L., Knecht, S., Angeli, C., and Reiher, M. (2017). Multireference perturbation theory with Cholesky decomposition for the density matrix renormalization group. *J. Chem. Theory Comput.* 13 (2): 451–459.

60 Zgid, D., Ghosh, D., Neuscamman, E., and Chan, G.K.-L. (2009). A study of cumulant approximations to *n*-electron valence multireference perturbation theory. *J. Chem. Phys.* 130 (19): 194107.

61 Sharma, S. and Chan, G.K.-L. (2014). Communication: a flexible multi-reference perturbation theory by minimizing the Hylleraas functional with matrix product states. *J. Chem. Phys.* 141 (11): 111101.

62 Sharma, S., Jeanmairet, G., and Alavi, A. (2016). Quasi-degenerate perturbation theory using matrix product states. *J. Chem. Phys.* 144 (3): 034103.

63 Sokolov, A.Y. and Chan, G.K.-L. (2016). A time-dependent formulation of multi-reference perturbation theory. *J. Chem. Phys.* 144 (6): 064102.

64 Roemelt, M., Guo, S., and Chan, G.K.-L. (2016). A projected approximation to strongly contracted *n*-electron valence perturbation theory for DMRG wave functions. *J. Chem. Phys.* 144 (20): 204113.

65 Phung, Q.M., Wouters, S., and Pierloot, K. (2016). Cumulant approximated second-order perturbation theory based on the density matrix renormalization group for transition metal complexes: a benchmark study. *J. Chem. Theory Comput.* 12 (9): 4352–4361.

66 Sokolov, A.Y., Guo, S., Ronca, E., and Chan, G.K.-L. (2017). Time-dependent *n*-electron valence perturbation theory with matrix product state reference wave functions for large active spaces and basis sets: applications to the chromium dimer and all-trans polyenes. *J. Chem. Phys.* 146 (24): 244102.

67 Lindh, R. and Roos, B.O. (1989). A theoretical study of the diffuseness of the $v(^1b_{1u})$ state of planar ethylene. *Int. J. Quantum Chem.* 35 (6): 813–825.

68 Lindgren, I. (1974). The Rayleigh–Schrodinger perturbation and the linked-diagram theorem for a multi-configurational model space. *J. Phys. B: At. Mol. Phys.* 7 (18): 2441–2470.

69 Lindgren, I. (1978). A coupled-cluster approach to the many-body perturbation theory for open-shell systems. *Int. J. Quantum Chem.* 14 (S12): 33–58.

70 Pittner, J. (2003). Continuous transition between Brillouin–Wigner and Rayleigh–Schrödinger perturbation theory, generalized Bloch equation, and Hilbert space multireference coupled cluster. *J. Chem. Phys.* 118 (24): 10876–10889.

71 Nakano, H. (1993). Quasidegenerate perturbation theory with multiconfigurational self-consistent-field reference functions. *J. Chem. Phys.* 99 (10): 7983–7992.

72 Nakano, H., Uchiyama, R., and Hirao, K. (2002). Quasi-degenerate perturbation theory with *general* multiconfiguration self-consistent field reference functions. *J. Comput. Chem.* 23 (12): 1166–1175.

73 Angeli, C., Borini, S., Cestari, M., and Cimiraglia, R. (2004). A quasidegenerate formulation of the second order *n*-electron valence state perturbation theory approach. *J. Chem. Phys.* 121 (9): 4043–4049.

74 Finley, J., Malmqvist, P.-Å., and Björn, O. (1998). Roos, and Luis Serrano-Andrés. The multi-state CASPT2 method. *Chem. Phys. Lett.* 288 (2–4): 299–306.

75 Zaitsevskii, A. and Malrieu, J.-P. (1995). Multi-partitioning quasidegenerate perturbation theory. A new approach to multireference Møller–Plesset perturbation theory. *Chem. Phys. Lett.* 233 (5–6): 597–604.

76 Serrano-Andrés, L., Merchán, M., and Lindh, R. (2005). Computation of conical intersections by using perturbation techniques. *J. Chem. Phys.* 122 (10): 104107.

77 Granovsky, A.A. (2011). Extended multi-configuration quasi-degenerate perturbation theory: the new approach to multi-state multi-reference perturbation theory. *J. Chem. Phys.* 134 (21): 214113.

78 Shiozaki, T., Győrffy, W., Celani, P., and Werner, H.-J. (2011). Communication: extended multi-state complete active space second-order perturbation theory: energy and nuclear gradients. *J. Chem. Phys.* 135 (8): 081106.

79 Andersson, K. (1992). *Multiconfigurational perturbation theory*. PhD thesis. Lund University, Lund, Sweden. https://urn.kb.se/resolve?urn=urn:nbn:se:kau:diva-21628.

Appendix

The following is an almost verbatim copy of appendix A in K. Andersson's thesis [79], reproduced here with her kind permission. Some adaptation has been applied in order to maintain consistency with the notation used in the rest of this chapter. Where different this will be pointed out.

This appendix is a collection of all matrix elements and vector components used in the first-order equation (see for example eqs. 10.24 and 10.108). We note that the orbital basis used in the appendix is the one in which the diagonal blocks of the Fock matrix (inactive, active and virtual), have been diagonalized. To emphasize this the Fock matrix elements are denoted f' so that there is a distinction with respect to a Fock matrix in a general orbital representation, f.

The following additional definitions and notations have been used (note that the two-particle density-matrix elements, Γ_{pqrs}, carry a factor $\frac{1}{2}$ in the work of Andersson, while here we stand by the definition in Chapter 1, section 1.6):

$$\hat{e}_{ut} = 2\delta_{ut} - \hat{E}_{ut} \ ,$$

$$d_{ut} = \langle \Psi^{(0)} | \hat{e}_{ut} | \Psi^{(0)} \rangle \ , \text{and}$$

$$\gamma_{utxv} = \langle \Psi^{(0)} | \hat{e}_{ut} \hat{e}_{xv} - \delta_{uv} \hat{e}_{xt} | \Psi^{(0)} \rangle \ .$$

The index notation specified in Chapter 1, section 1.5, will be used strictly in this appendix except in the following formula:

$$f^{\mathrm{I}}_{pq} = h_{pq} + \sum_k [2(pq|kk) - (pk|kq)] \ ,$$

where k runs over both the *frozen* and inactive subspaces, i.e., over all doubly-occupied orbitals.

In the evaluation of the matrix elements and vector components the following commutation relations have been useful:

$$[\hat{E}_{pq}, \hat{E}_{rs}] = \delta_{rq}\hat{E}_{ps} - \delta_{ps}\hat{E}_{rq} \ ,$$
$$[\hat{e}_{pq}, \hat{e}_{rs}] = \delta_{ps}\hat{e}_{rq} - \delta_{rq}\hat{e}_{ps} \ .$$

Matrix elements and vector components, with respect to combinations stemming from interactions between states of the same class within the first-order interacting space, will be presented in the following section. Corresponding matrix elements between states of different classes of the first-order interacting space will be presented subsequently. We note also that we will use the following abbreviation: $|\Psi^{(0)}\rangle = |0\rangle$ and similarly for the bra.

The Diagonal Blocks

The following elements are given in this section:

$$\langle\sigma|\hat{H}_0 - E^{(0)}|\tau\rangle \ , \ \langle\sigma|V|0\rangle \ , \ |\sigma\rangle, |\tau\rangle \in V_X \ ,$$

where $\hat{H}_0 = \sum_{pq} f'_{pq}\hat{E}_{pq}$ and $X \in \{A, \ldots, H\}$. We note that, $f'_{pq} = \delta_{pq}\epsilon_p$, if both indices, p and q, belong to the same orbital class, i.e., inactive, active, and virtual. Non-zero elements are found in the four inactive–active and active–virtual blocks.

Internal excitations: $X \in \{A, B\}$

A	***B*** $\quad i \geq j, k \geq l, t \geq u, x \geq y$
$\|ituv\rangle = \hat{E}_{ti}\hat{E}_{uv}\|0\rangle$	$\|ijtu\rangle^{\pm} = [\hat{E}_{ij}\hat{E}_{ui} \pm \hat{E}_{ti}\hat{E}_{uj}]\|0\rangle$
$\langle ituv\|\hat{H}_0 - E^{(0)}\|kxyz\rangle$	${}^{\pm}\langle ijtu\|\hat{H}_0 - E^{(0)}\|klxy\rangle^{\pm}$
$\quad = \delta_{ik}[\alpha^i S^A_{tuv,xyz} + B_{tuv,xyz}]$	$\quad = \delta_{ik}\delta_{jl}[1 \pm \delta_{ij}][\alpha^{ij} S^{\pm}_{tu,xy} + B^{\pm}_{tu,xy}]$
$\langle ituv\|V\|0\rangle = V^i_{tuv}$	${}^{\pm}\langle ijtu\|V\|0\rangle = \frac{1}{2}[1 \pm \delta_{ij}](\boldsymbol{S}^{\pm}\boldsymbol{W}^{ij,\pm})_{tu}$
$\alpha^i = -\epsilon_i - \sum_w \epsilon_w D_{ww}$	$\alpha^{ij} = -\epsilon_i - \epsilon_j + \sum_w \epsilon_w d_{ww}$
$S^A_{tuv,xyz} = \langle 0\|\hat{E}_{vu}\hat{e}_{xt}\hat{E}_{yz}\|0\rangle$	$S^{\pm}_{tu,xy} = S_{tu,xy} \pm S_{tu,yx}, \ S_{tu,xy} = 2\gamma_{xtyu}$
$B_{tuv,xyz} = \epsilon_x\langle 0\|\hat{E}_{vu}\hat{e}_{xt}\hat{E}_{yz}\|0\rangle$	$B^{\pm}_{tu,xy} = B_{tu,xy} \pm B_{tu,yx}$
$\quad + \sum_w \epsilon_w\langle 0\|\hat{E}_{vu}\hat{e}_{xt}\hat{E}_{ww}\hat{E}_{yz}\|0\rangle$	$B_{tu,xy} = [\epsilon_t + \epsilon_y][2\gamma_{xtyu} + \delta_{ty}d_{xu}]$
$V^i_{tuv} = 2f^{\mathrm{I}}_{ti}D_{vu} - f^{\mathrm{I}}_{ui}D_{vt} - \sum_w \Gamma_{vuww} f^{\mathrm{I}}_{wi}$	$\quad + 2\sum_w \epsilon_w[\delta_{ty}\gamma_{xuww} - \langle 0\|\hat{e}_{yu}\hat{e}_{ww}\hat{e}_{xt}\|0\rangle]$
$\quad + \sum_{wxy}\langle 0\|\hat{E}_{vu}\hat{e}_{wt}\hat{E}_{xy}\|0\rangle(wi\|xy)$	$W^{ij,\pm}_{tu} = \frac{1}{1+\delta_{ij}}\left[1 - \frac{1}{2}\delta_{tu}\right][(ti\|uj) \pm (ui\|tj)]$

Semi-internal excitations: $X \in \{C, D, E\}$

C

$|atuv\rangle = \hat{E}_{at}\hat{E}_{uv}|0\rangle$

$\langle atuv|\hat{H}_0 - E^{(0)}|cxyz\rangle$
$= \delta_{ac}[\alpha^a S^C_{tuv,xyz} + B_{tuv,xyz}]$

$\langle atuv|V|0\rangle = V^a_{tuv}$

$\alpha^a = \epsilon_a - \sum_w \epsilon_w D_{ww}$

$S^C_{tuv,xyz} = \langle 0|\hat{E}_{vu}\hat{E}_{tx}\hat{E}_{yz}|0\rangle$

$B_{tuv,xyz} = -\epsilon_x\langle 0|\hat{E}_{vu}\hat{E}_{tx}\hat{E}_{yz}|0\rangle$
$+\sum_w \epsilon_w \langle 0|\hat{E}_{vu}\hat{E}_{tx}\hat{E}_{ww}\hat{E}_{yz}|0\rangle$

$V^a_{tuv} = \sum_{wxy}\langle 0|\hat{E}_{vu}\hat{E}_{tw}\hat{E}_{xy}|0\rangle(aw|xy)$
$+\sum_w[\Gamma_{vutw} + \delta_{ut}D_{vw}]\left[f^{\mathrm{I}}_{aw} - \sum_y (ay|yw)\right]$

D

$|iatu\rangle^1 = \hat{E}_{ai}\hat{E}_{tu}|0\rangle$, $|iatu\rangle^2 = \hat{E}_{ti}\hat{E}_{au}|0\rangle$

${}^Q\langle iatu|\hat{H}_0 - E^{(0)}|kcxy\rangle^P$
$= \delta_{ac}\delta_{ik}[\alpha^{ia} S^{Q,P}_{tu,xy} + B^{Q,P}_{tu,xy}]$

${}^Q\langle iatu|V|0\rangle = V^{ia,Q}_{tu}$ $\quad P, Q \in \{1, 2\}$

$\alpha^{ia} = \epsilon_a - \epsilon_i - \sum_w \epsilon_w D_{ww}$

$S = \begin{bmatrix} S^{1,1} & S^{1,2} \\ S^{2,1} & S^{2,2} \end{bmatrix}, B = \begin{bmatrix} B^{1,1} & B^{1,2} \\ B^{2,1} & B^{2,2} \end{bmatrix}$

$S^{1,1}_{tu,xy} = -2S^{2,1}_{tu,xy} = 2[\Gamma_{utxy} + \delta_{tx}D_{uy}]$

$S^{2,2}_{tu,xy} = -[\Gamma_{xtuy} - 2\delta_{tx}D_{uy}]$

$B^{1,1}_{tu,xy} = -2B^{2,1}_{tu,xy}$
$= 2\sum_w \epsilon_w \langle 0|\hat{E}_{ut}\hat{E}_{ww}\hat{E}_{xy}|0\rangle$

$B^{2,2}_{tu,xy} = [\epsilon_u - \epsilon_x]\Gamma_{xtuy} + \epsilon_u\delta_{tu}D_{xy}$
$+2\epsilon_x\delta_{tx}D_{uy} + \sum_w \epsilon_w[2\delta_{tx}\Gamma_{wwuy}$
$+\delta_{tu}\Gamma_{wuxy} - \langle 0|\hat{E}_{xt}\hat{E}_{ww}\hat{E}_{uy}|0\rangle]$

$\boldsymbol{V}^{ia} = f^{\mathrm{I}}_{ai}\boldsymbol{U} + \boldsymbol{S}\boldsymbol{W}^{ia}$

$\boldsymbol{V}^{ia} = \begin{bmatrix} V^{ia,1} \\ V^{ia,2} \end{bmatrix}, \boldsymbol{U} = \begin{bmatrix} U^1 \\ U^2 \end{bmatrix}, \boldsymbol{W}^{ia} = \begin{bmatrix} W^{ia,1} \\ W^{ia,2} \end{bmatrix}$

$U^1_{tu} = 2D_{ut}$, $U^2_{tu} = -D_{ut}$

$W^{ia,1}_{tu} = (ai|tu)$, $W^{ia,2}_{tu} = (ti|au)$

E $\quad i \geq j, k \geq l$

$|ijat\rangle^{\pm} = [\hat{E}_{tj}\hat{E}_{ai} \pm \hat{E}_{ti}\hat{E}_{aj}]|0\rangle$

${}^{\pm}\langle ijat|\hat{H}_0 - E^{(0)}|klcx\rangle^{\pm}$
$= 2[2 \mp 1]\delta_{ac}\delta_{ik}\delta_{jl}[1 \pm \delta_{ij}][\alpha^{ija} S_{tx} + B_{tx}]$

${}^{\pm}\langle ijat|V|0\rangle = [2 \mp 1][1 \pm \delta_{ij}](\boldsymbol{S}\boldsymbol{W}^{ija,\pm})_t$

$\alpha^{ija} = \epsilon_a - \epsilon_i - \epsilon_j + \sum_w \epsilon_w d_{ww}$

$S_{tx} = d_{xt}$

$B_{tx} = -\sum_w \epsilon_w \gamma_{wwxt}$

$W^{ija,\pm}_t = \frac{1}{1+\delta_{ij}}[(ai|tj) \pm (aj|ti)]$

F $\quad a \geq b, c \geq d, t \geq u, x \geq y$

$|abtu\rangle^{\pm} = [\hat{E}_{bt}\hat{E}_{au} \pm \hat{E}_{at}\hat{E}_{bu}]|0\rangle$

${}^{\pm}\langle abtu|\hat{H}_0 - E^{(0)}|cdxy\rangle^{\pm}$
$= \delta_{ac}\delta_{bd}[1 \pm \delta_{ab}][\alpha^{ab} S^{\pm}_{tu,xy} + B^{\pm}_{tu,xy}]$

${}^{\pm}\langle abtu|V|0\rangle = \frac{1}{2}[1 \pm \delta_{ab}](\boldsymbol{S}^{\pm}\boldsymbol{W}^{ab,\pm})_{tu}$

$\alpha^{ab} = \epsilon_a + \epsilon_b - \sum_w \epsilon_w D_{ww}$

$S^{\pm}_{tu,xy} = S_{tu,xy} \pm S_{tu,yx}$, $S_{tu,xy} = 2\Gamma_{txuy}$

$B^{\pm}_{tu,xy} = B_{tu,xy} \pm B_{tu,yx}$

$B_{tu,xy} = -[\epsilon_t + \epsilon_y][2\Gamma_{txuy} + \delta_{ty}D_{ux}]$
$-2\sum_w \epsilon_w[\delta_{ty}\Gamma_{uxww} - \langle 0|\hat{E}_{uy}\hat{E}_{ww}\hat{E}_{tx}|0\rangle]$

$W^{ab,\pm}_{tu} = \frac{1}{1+\delta_{ab}}\left[1 - \frac{1}{2}\delta_{tu}\right][(au|bt) \pm (at|bu)]$

G $\quad a \geq b, c \geq d$

$|iabt\rangle^{\pm} = [\hat{E}_{bi}\hat{E}_{at} \pm \hat{E}_{ai}\hat{E}_{bt}]|0\rangle$

${}^{\pm}\langle iabt|\hat{H}_0 - E^{(0)}|kcdx\rangle^{\pm}$
$= 2[2 \mp 1]\delta_{ik}\delta_{ac}\delta_{bd}[1 \pm \delta_{ab}][\alpha^{iab} S_{tx} + B_{tx}]$

${}^{\pm}\langle iabt|V|0\rangle = [2 \mp 1][1 \pm \delta_{ab}](\boldsymbol{S}\boldsymbol{W}^{iab,\pm})_t$

$\alpha^{iab} = \epsilon_a + \epsilon_b - \epsilon_i - \sum_w \epsilon_w D_{ww}$

$S_{tx} = D_{tx}$

$B_{tx} = \sum_w \epsilon_w \Gamma_{wwtx}$

$W^{iab,\pm}_t = \frac{1}{1+\delta_{ab}}[(at|bi) \pm (bt|ai)]$

External excitations: $X \in \{F, G, H\}$

H $\quad i \geq j, k \geq l, a \geq b, c \geq d$

$|ijab\rangle^{\pm} = [\hat{E}_{bj}\hat{E}_{ai} \pm \hat{E}_{aj}\hat{E}_{bi}]|0\rangle$

${}^{\pm}\langle ijab|\hat{H}_0 - E^{(0)}|klcd\rangle^{\pm} = 4[2 \mp 1]\delta_{ik}\delta_{jl}\delta_{ac}\delta_{bd}[1 \pm \delta_{ij}][1 \pm \delta_{ab}][\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j]$

${}^{\pm}\langle ijab|V|0\rangle = 2[2 \mp 1][(ai|bj) \pm (aj|bi)]$

The Nondiagonal Blocks

The following elements are given in this section:

$$\langle\sigma|\hat{H}_0|\tau\rangle\ ,\ |\sigma\rangle \in V_X\ ,\ |\tau\rangle \in V_Y\ ,\ X \neq Y \in \{A,\ldots,H\}\ .$$

AB $k \geq l, x \geq y$

$$\langle ituv|\hat{H}_0|klxy\rangle^{\pm} = \sum_w [\delta_{il} f'_{wk} \pm \delta_{ik} f'_{wl}]\{-[S^A_{tuv,xyw} \pm S^A_{tuv,yxw}]$$
$$+[2 \mp 1][\delta_{yw}(2\delta_{xt} D_{vu} - \delta_{ux} D_{vt} - \Gamma_{vuxt}) \pm \delta_{xw}(2\delta_{yt} D_{vu} - \delta_{uy} D_{vt} - \Gamma_{vuyt})]\}$$

CF $c \geq d, x \geq y$

$$\langle atuv|\hat{H}_0|cdxy\rangle^{\pm} = \sum_w [\delta_{ad} f'_{wc} \pm \delta_{ac} f'_{wd}]$$
$$\times\{[S^C_{tuv,xwy} - \delta_{ux}(\Gamma_{vuty} + \delta_{ut} D_{vy})] \pm [S^C_{tuv,ywx} - \delta_{uy}(\Gamma_{vutx} - \delta_{ut} D_{vx})]\}$$

AD

$$\langle ituv|\hat{H}_0|kcxy\rangle^1 = \delta_{ik} \sum_w f'_{wc} S^A_{tuv,wxy}$$
$$\langle ituv|\hat{H}_0|kcxy\rangle^2 = \delta_{ik} \sum_w f'_{wc} S^A_{tuv,xwy}$$

CD

$$\langle atuv|\hat{H}_0|kcxy\rangle^1 = -\delta_{ac} \sum_w f'_{wk} S^C_{tuv,wxy}$$
$$\langle atuv|\hat{H}_0|kcxy\rangle^2 = -\delta_{ac} \sum_w f'_{wk} [S^C_{tuv,yxw}$$
$$-2\delta_{ux}(\Gamma_{vuty} + \delta_{ut} D_{vy}) - \delta_{xy}(\Gamma_{vutw} + \delta_{ut} D_{vw})]$$

BE $i \geq j, k \geq l, t \geq u$

$${}^{\pm}\langle ijtu|\hat{H}_0|klcx\rangle^{\pm} = 2\delta_{ik}\delta_{jl}[1 \pm \delta_{ij}] \sum_w f'_{wc} [\gamma_{xtwu} \pm \gamma_{xuwt}]$$

DE $k \geq l$

$${}^1\langle iatu|\hat{H}_0|klcx\rangle^{\pm} = -[2 \mp 1]\delta_{ac} \sum_w [\delta_{ik} f'_{wl} \pm \delta_{il} f'_{wk}][\Gamma_{utxw} + \delta_{tx} D_{uw} - 2\delta_{xw} D_{ut}]$$

$${}^2\langle iatu|\hat{H}_0|klcx\rangle^{\pm} = \delta_{ac} \sum_w [\delta_{il} f'_{wk} \pm \delta_{ik} f'_{wl}]$$
$$\times\{[\Gamma_{uwxt} + \delta_{xw} D_{ut} - 2\delta_{tx} D_{uw}] \pm [\Gamma_{utxw} + \delta_{tx} D_{uw} - 2\delta_{xw} D_{ut}]\}$$

FG $a \geq b, c \geq d, t \geq u$

$${}^{\pm}\langle abtu|\hat{H}_0|kcdx\rangle^{\pm} = -2\delta_{ac}\delta_{bd}[1 \pm \delta_{ab}] \sum_w f'_{wk} [\Gamma_{uxtw} \pm \Gamma_{uwtx}]$$

DG $c \geq d$

$${}^1\langle iatu|\hat{H}_0|kcdx\rangle^{\pm} = [2 \mp 1]\delta_{ik} \sum_w [\delta_{ad} f'_{wc} \pm \delta_{ac} f'_{wd}][\Gamma_{utwx} + \delta_{tw} D_{ux}]$$

$${}^2\langle iatu|\hat{H}_0|kcdx\rangle^{\pm} = \delta_{ik} \sum_w [\delta_{ad} f'_{wc} \pm \delta_{ac} f'_{wd}]\{[-\Gamma_{utwx} - \delta_{tw} D_{ux}] \pm [-\Gamma_{uxwt} + 2\delta_{tw} D_{ux}]\}$$

EH $i \geq j, k \geq l, c \geq d$

$${}^{\pm}\langle ijat|\hat{H}_0|klcd\rangle^{\pm} = 2[2 \mp 1]\delta_{ik}\delta_{jl}[1 \pm \delta_{ij}] \sum_w [\delta_{ac} f'_{wd} \pm \delta_{ad} f'_{wc}] d_{wt}$$

GH $a \geq b, k \geq l, c \geq d$

$${}^{\pm}\langle iabt|\hat{H}_0|klcd\rangle^{\pm} = -2[2 \mp 1]\delta_{ac}\delta_{bd}[1 \pm \delta_{ab}] \sum_w [\delta_{il} f'_{wk} \pm \delta_{ik} f'_{wl}] D_{tw}$$

The remaining upper-triangular blocks are equal to zero and the lower-triangular blocks are given by symmetry, since $\hat{H}_0$ is a Hermitian operator.

Part II

Nuclear Dynamics

11

Exact Quantum Dynamics (Wave Packets) in Reduced Dimensionality

Sebastian Reiter, Daniel Keefer, and Regina de Vivie-Riedle

Department Chemie, Ludwig-Maximilians-Universität München, Butenandtstr. 11, 81377 München, Germany

Abstract

The present chapter outlines basic and advanced concepts of grid-based quantum dynamics for molecular systems. Simulations within this framework are used to investigate the time-evolution of a molecular quantum system during a physical or chemical process of interest. The goal is to give the reader a concise introduction to wave packet simulations and strategies to reduce complex systems to few coordinates. Strengths and limitations are discussed using applied examples. A comprehensive flowchart on how to set up a wave packet simulation is given. Being principally exact, a solution of the time-dependent Schrödinger equation for the nuclear dynamics is only feasible in few dimensions due to exponential computational cost. This means the most crucial step is to find a representation of the molecular process using only few important coordinates. The concept of reactive coordinates is introduced, being determined either by chemical intuition or the adaptation of machine learning techniques. Using a certain reduced-dimensional representation, all terms within the molecular Hamiltonian are discussed along with the means to obtain them. A special focus lies on the kinetic energy operator, where the G-Matrix formalism is introduced as a very general scheme to transform it from cartesian coordinates to an arbitrary set of linear or non-linear reactive coordinates. Another focus lies on the evaluation of non-adiabatic coupling matrix elements and their implementation within the reduced-dimensional quantum dynamical framework. This allows for the simulation of wave packets passing through conical intersections, a feature determining the outcome of virtually all fast photochemical processes. Applied examples using the introduced concepts are illustrated.

11.1 Introduction

In nuclear quantum dynamics, the time evolution of atomic nuclei is simulated by solving the time-dependent Schrödinger equation (TDSE). Within the principally exact TDSE framework, the molecular motion is represented by wave packets on electronic potential energy surfaces (PESs). This allows for the description of quantum effects like coherence phenomena, tunneling processes, branching at conical intersections or the interaction with light (e.g., ultrashort laser pulses). The molecular wave packet is usually discretized on a spatial grid, where an exact (i.e., full-dimensional) wave packet representation of a non-linear molecule with N atoms includes $3N$-6 intramolecular degrees of freedom. Thus, a small organic molecule like benzene already contains 30 internal degrees of freedom. Assuming a required amount of 128 grid points in each

Quantum Chemistry and Dynamics of Excited States: Methods and Applications,
First Edition. Edited by Leticia González and Roland Lindh.

dimension, a full-dimensional description of benzene would require $128^{30} = 1.6 \times 10^{63}$ grid points, which is very far from being realizable on modern computational architectures and known as the curse of dimensionality. However, especially in ultrafast chemical processes, the number of relevant degrees of freedom involved in a chemical reaction is often much smaller than 3*N*-6, and the dimensionality of the problem can be reduced. Specifically, the typical number of dimensions included in a simulation on explicit ab initio PESs is between one and three. For the previous example, this would reduce the number of grid points to a feasible number between $128^1 = 128$ and $128^3 = 2{,}097{,}152$ grid points. This reduction of dimensionality is the biggest approximation in grid-based quantum dynamics. Therefore, this chapter will outline criteria and methods to arrive at an optimal low-dimensional coordinate space for a given molecular process, along with applied simulation examples from the chemical and biological domain.

After a general introduction to quantum dynamics, we will begin with the most intuitive approach, which is the manual selection of coordinates by chemical intuition. Chemical motions like bond-cleavage or proton transfer can be described, for example, by taking normal modes such as changes in bond distances or angles as reactive coordinates and omitting all other degrees of freedom. This approach will be demonstrated using the example of the photoinduced ring opening of cyclohexadiene and can produce meaningful results in straightforward cases. In more complex cases, it is also possible to construct reactive coordinates by combining several normal modes in a single displacement vector, which will be exemplified by the ultrafast photorelaxation of the RNA-nucleobase uracil. These approaches rely on the chemical intuition of the researcher and/or extensive preliminary work. This can become very challenging for complex processes with many contributing modes. It is therefore desirable to automate this process and leave the search for objectively well-suited coordinate spaces to machines. Following on from this idea, three (semi-)automatic techniques to obtain meaningful coordinates for quantum dynamical simulations will also be described in this chapter. This comprises the use of machine-learning methodologies, which can even enable the construction of reasonable non-linear coordinate spaces. In addition, guidelines will be given to set up the complete Hamiltonian for quantum dynamical simulations in arbitrary reduced coordinates. This involves the computation and use of PESs and non-adiabatic couplings for the potential energy operator, as well as a general transformation protocol for the kinetic energy operator from Cartesian coordinates x to arbitrary coordinates q.

11.2 Fundamentals of Molecular Quantum Dynamics

Due to the complexity of quantum dynamical simulations, analytical solutions are rarely possible and numerical propagation schemes need to be applied. The calculations can be performed based either on a grid or on an eigenstate representation [1]. In the latter case, the eigenfunctions and eigenvalues have to be known. Both approaches differ in the choice of the basis functions and can easily be transferred to each other. In the following, the basic tools to solve the TDSE for complex molecules will be summarized.

11.2.1 Wave Packet Dynamics

The time evolution of a molecular system is governed by the TDSE:

$$i\hbar \frac{\partial}{\partial t} \Psi_{\mathrm{mol}}(t) = \hat{H}_{\mathrm{tot}} \Psi_{\mathrm{mol}}(t) \tag{11.1}$$

with the Hamiltonian $\hat{H}_{\text{tot}}$, comprising the molecular Hamiltonian $\hat{H}_0$ and possible external perturbations, and the system's wave function $\Psi_{\text{mol}}(t)$. The wave function can formally be separated into nuclear $\Psi_{\text{nuc}}(R, t)$ and electronic $\Phi_{\text{el}}(r, t; R)$ wave functions, which depend parametrically on the nuclear coordinates R:

$$\Psi_{\text{mol}}(t) = \Psi_{\text{nuc}}(R, t)\Phi_{\text{el}}(r, t; R) \tag{11.2}$$

In this chapter, we focus on the intramolecular nuclear dynamics where only the nuclear wave function is propagated in time. In addition, the Born–Oppenheimer approximation is used, which allows for a separation of electronic and nuclear motion. The electronic Schrödinger equation can be solved with standard quantum chemical program packages for different nuclear configurations R to obtain the respective PES. The intramolecular motion of the nuclei is thus described by the evolution of the nuclear wave function $\Psi_{\text{nuc}}(R, t)$ on a pre-calculated PES. From now on the subscript *nuc* as well as the R-dependence will be omitted for the nuclear wave function, which will be denoted by $\Psi(t)$, except if the explicit form is needed for clarification. For a time-independent Hamiltonian $\hat{H}_0$ including the kinetic energy operator $\hat{T}$ and the potential energy operator $\hat{V}$ of the system, the nuclear Schrödinger equation now reads:

$$i\hbar\frac{\partial}{\partial t}\Psi(t) = \hat{H}_0\Psi(t). \tag{11.3}$$

Integrating the time-dependent nuclear Eq. (11.3) determines the equations of motion as the action of a propagator $U(t, t_0)$ on the nuclear wave function:

$$\Psi(t) = U(t, t_0)\Psi(t_0) = e^{-\hat{H}_0(t-t_0)}\Psi(t_0). \tag{11.4}$$

Due to the propagator $U(t, t_0)$ the nuclear wave function evolves from the initial time t_0 to the final time t under the influence of its Hamiltonian. For the case where the eigenvalue problem of $\hat{H}_0$ is solved, the propagator can be written in the basis of the vibrational eigenfunctions ψ_i

$$U(t, t_0) = \sum_i e^{-iE_i(t-t_0)/\hbar}|\psi_i\rangle\langle\psi_i|, \tag{11.5}$$

with the corresponding vibrational eigenenergies E_i. Now we can answer the question: When does the time evolution of a nuclear wave function show up? The observable is the probability density given by the square amplitude of the wave function. In the case of a single vibrational wave function ψ_{E_j} with eigenvalue E_j (which is called a special solution in David Tannor's book [1]) no temporal evolution is observable:

$$|\Psi_j(R, t)|^2 = (\psi_{E_j}(R)e^{-iE_j\,t/\hbar})^*(\psi_{E_j}(R)e^{-iE_j\,t/\hbar}) = |\Psi_j(R)|^2. \tag{11.6}$$

The more general solution for the time-dependent nuclear wave function, also referred to as a wave packet, is a linear combination of special solutions:

$$\Psi_{\text{wp}}(R, t) = \sum_i c_i\psi_i e^{-\frac{i}{\hbar}E_i t}. \tag{11.7}$$

For the easiest case, i.e., a two-state superposition, we now observe a temporal change in the probability density, meaning that the wave packet moves around in the coordinate space:

$$|\Psi_{\text{wp}}(R, t)|^2 = |c_1|^2|\psi_{E_1}(R)|^2 + |c_2|^2|\psi_{E_2}(R)|^2 + 2\text{Re}\left\{c_1^* c_2\psi_{E_1}^*(R)\psi_{E_2}(R)e^{-\frac{i}{\hbar}(E_2-E_1)t}\right\}. \tag{11.8}$$

Note that the third term, called the interference term between the two special solutions, contains all the time-dependence and coherence information of the wave packet.

11.2.2 Time-Propagator Schemes

Molecular chemical processes typically include bond breaking and formation. Thus, not only bound vibrational states but also continuum states are involved, which makes the pre-calculation of the respective eigenfunction, i.e., the eigenstate approach, more elaborate and the grid approach more favorable. One of the widely used grid methods is the pseudospectral method which operates on a basis of unit delta functions located at each of the grid points in the nuclear coordinate space. In this spatial representation, the PES and all possible coupling elements are simply given by their value at the grid point and the wave functions by their amplitudes. As a special case of pseudospectral methods, the Fourier method operates on a grid of evenly spaced points. It has two implementations: the Fourier Grid Hamiltonian method [2], which involves the construction of the Hamilton matrix in the pseudospectral Fourier basis to evaluate the bound state eigenvalues and eigenstates and the Dynamic Fourier Method [3], in which the action of the Hamiltonian on the wave function $H\Psi(t)$ is calculated directly.

The Dynamic Fourier Method requires explicit propagator schemes, all relying on time discretization. The total propagator for the time interval $[t_0,t]$ is written as a product function over small time intervals Δt with $N\Delta t = t$

$$U(t,t_0) = e^{-i\hat{H}_0(t-t_0)/\hbar} = \underbrace{e^{-i\hat{H}_0\Delta t/\hbar}\cdot e^{-i\hat{H}_0\Delta t/\hbar}\cdots e^{-i\hat{H}_0\Delta t/\hbar}}_{N \text{ times}} \tag{11.9}$$

This formulation also allows the time-dependent perturbation to be included in the Hamiltonian, such as an external electric field $\varepsilon(t)$ interacting with the molecular system. The light field

$$\varepsilon(t) = \varepsilon_0 exp\left[\left[-\left(\frac{t}{\mathrm{FWHM}/\sqrt{2ln2}}\right)^2\right] cos(\omega(t)t)\right] \tag{11.10}$$

is described in the dipole approximation with the field amplitude ε_0, the frequency ω, and a Gaussian shaped field envelope characterized by the full width half maximum (FWHM) of the light pulse. The coupling to the molecule is mediated by the dipole moment μ and the total Hamiltonian becomes:

$$\hat{H}_{\mathrm{tot}} = \hat{H}_0 - \mu\varepsilon(t). \tag{11.11}$$

For any time-dependent Hamiltonian $\hat{H}_{\mathrm{tot}}$, the propagation has to be performed in sufficiently small time steps, so that the perturbation can be regarded as constant during the time interval Δt. From Eq. (11.9) it becomes clear that the Hamiltonian has to be applied to the wave function in every time step, requiring a fast operation. The breakthrough came with the fast Fourier transform technique for the evaluation of the kinetic part of the Hamiltonian [4, 5]. This method takes advantage of the fact that the momentum operator and its square are non-local in the coordinate representation but local in the momentum representation, where their action can be evaluated by simple multiplication. For the complete operation $\hat{H}\Psi(t)$ the wave function is Fourier-transformed to the momentum space before the second derivative operator of the kinetic part is applied:

$$\frac{\partial^2}{\partial x^2}\Psi(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\frac{\partial^2}{\partial x^2}[\Psi(k)e^{ikx}]dk = \frac{1}{2\pi}\int_{-\infty}^{+\infty} -k^2\Psi(k)e^{ikx}dk. \tag{11.12}$$

As the potential energy operator is local in the coordinate representation, the wave function has to be transformed back to the real space. In the context of quantum dynamics, the sequence

$$\Psi(x) \xrightarrow{IFT} \Psi(k) \rightarrow -k^2\Psi(k) \xrightarrow{FT} \frac{\partial^2}{\partial x^2}\Psi(x) \tag{11.13}$$

is applied in each time step Δt .

The grid space representation of the Fourier method requires the discretized form of the Fourier transform (DFT) and its inverse (IDFT)

$$\begin{aligned} \text{DFT:} \quad F(k_j) &= \frac{1}{N} \sum_{n=-\frac{N}{2}+1}^{\frac{N}{2}} f(x_n) e^{-inj/N} \\ \text{IDFT:} \quad f(x_n) &= \sum_{j=-\frac{N}{2}+1}^{\frac{N}{2}} F(k_j) e^{inj/N} \end{aligned} \tag{11.14}$$

The total range $A{=}N\Delta x$ in coordinate space is defined by the number N of grid points and the uniform spacing Δx between the grid points. The index n defines the grid points in coordinate space ($x_n{=}n\Delta x$) and the corresponding index j in momentum space ($k_j{=}j\Delta k$). The grid size and spacing selected in the coordinate space determines the grid size in momentum space, whose central point is taken as $k_0{=}0$. Thus, A defines the longest wavelength and therefore the smallest frequency (k-value) in the momentum space $\Delta k{=}2\pi/N\Delta x$. To discretize the wave function without loss of information one has to follow Shannon's sampling theorem, which also directly links the spacing Δx with the largest representable momentum via $\Delta x \leq \frac{\pi\hbar}{p_{max}} = \frac{\pi}{k_{max}}$.

The numerical evaluation of the time propagator $e^{-i\hat{H}_{\text{tot}}\Delta t}\psi(t)$ can be performed efficiently with different techniques. A comparison of the various propagation schemes can be found in [6]. The easiest to implement is the second order difference (SOD) [6] method

$$e^{-\frac{i}{\hbar}\hat{H}_{\text{tot}}\Delta t} \approx 1 - \frac{i}{\hbar}\Delta t \hat{H}_{\text{tot}}, \tag{11.15}$$

in which the exponential functional is expanded in a Taylor series. Its further advantage is that there are no special requirements on the structure of the kinetic part of the Hamiltonian. A drawback is that the time steps chosen have to be very small, limiting the ability of this method for multi-dimensional long-time dynamics. A very efficient and accurate propagation scheme is the split operator (SPO) method introduced by Feit and Fleck [4]. The sum of the kinetic and potential energy operators in the exponent of the propagator is approximated by the product of the individual exponential operators:

$$e^{-\frac{i}{\hbar}\hat{H}_{\text{tot}}\Delta t} \approx e^{-\frac{i}{\hbar}\frac{\hat{T}}{2}\Delta t} e^{-\frac{i}{\hbar}\hat{V}\Delta t} e^{-\frac{i}{\hbar}\frac{\hat{T}}{2}\Delta t}. \tag{11.16}$$

In the symmetrized form of Eq. (11.16) the error introduced due to the non-commuting operators $\hat{T}$ and $\hat{V}$ can be kept to the order of $O(\Delta t)^3$. However, the SPO cannot be used when cross partial derivatives appear in the kinetic energy operator. In such situations the Chebychev polynomial expansion [7] can be applied as a fast and accurate method. Here, the propagator $U(t, t_0)$ is expressed as

$$e^{-\frac{i}{\hbar}\hat{H}_{tot}\Delta t} \equiv \sum_{n}^{N} a_n(t)\ \Phi_n(-i\hat{H}_{tot}), \tag{11.17}$$

where Φ_n are complex Chebychev polynomials, depending on the Hamiltonian and obeying the recursion relation:

$$\Phi_{n+1} = -2i\hat{H}_{tot}\Phi_n + \Phi_{n-1}. \tag{11.18}$$

The time-dependent expansion coefficients $a_n(t)$

$$\mathrm{a}_n(t) = 2J_n(t) \ \text{ and } \ a_0(t) = J_0(t) \tag{11.19}$$

are determined by Bessel functions $J_n(t)$. For the implementation, the argument of Φ_n has to be mapped onto the interval $[-i, i]$. Consequently, the eigenvalues of $\hat{H}_{\text{tot}}$ are scaled to the range $[-1, 1]$. The propagation is performed with the normalized Hamiltonian introducing a shift parameter to compensate for the normalization. The chosen order of the expansion N has to be large enough to ensure convergence of the series [7].

11.2.3 Excited State Wave Packet Dynamics

Fast dynamical processes are mostly started by photoexcitation into an excited electronic state with typical time scales in the femto- to picosecond range. The initial nuclear wave function is typically the lowest vibrational eigenfunction in the electronic groundstate. It can be evaluated, e.g., in the scope of the Dynamic Fourier Method by propagating the Schrödinger equation with a guess input function of Gaussian shape in imaginary time [8] or by applying the Fourier grid method as described by Marston et al. [2]. To initiate and follow their dynamics, femtosecond laser pulses are needed. Several PESs may be involved in these molecular processes and Eq. (11.1) changes to:

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix} a_1(t)\Psi_1(t) \\ \vdots \\ a_n(t)\Psi_n(t) \end{pmatrix} = \begin{pmatrix} H_{11} & \cdots & H_{1n} \\ \vdots & \ddots & \vdots \\ H_{n1} & \cdots & H_{nn} \end{pmatrix}\begin{pmatrix} a_1(t)\Psi_1(t) \\ \vdots \\ a_n(t)\Psi_n(t) \end{pmatrix}. \tag{11.20}$$

The nuclear wave function can now be distributed over several PESs. The distribution is given by the time-dependent expansion coefficients $a_n(t)$. The entries on the diagonal of the Hamiltonian matrix describe the uncoupled dynamics on the individual PES $H_{ii} = T_{ii} + V_{ii}(R)$ with the pre-calculated PES $V_{ii}(R)$ of state i, and the kinetic energy operation T_{ii}, which is equal for all states. Coupling between the different electronic states is described by the off-diagonal elements. They may reflect the interaction with the laser light, an intramolecular coupling like non-adiabatic and spin–orbit coupling, or all couplings simultaneously.

11.2.4 Surfaces and Coupling Elements in Reactive Coordinates

As our aim is to perform quantum dynamics in reduced dimensionality on highly accurate PESs, we need to find a way to express the PES and coupling elements in the reduced coordinates. The PES can be evaluated using any electronic structure level method appropriate for the system and the process. The respective structures are optimized in full dimensionality and then projected onto the reduced-dimensional space of reactive coordinates. After the projection, the Thin-Plate-Spline method [9] can be used to interpolate the PES. This allows its representation on arbitrarily spaced grids, which is necessary for wave packet propagations.

Ultrafast dynamics is often mediated by conical intersections and is described by non-adiabatic coupling elements [10]. They are neglected in the Born–Oppenheimer approximation but can be introduced as off-diagonal elements in Eq. (11.20). Two implementations are possible: the diabatic (see, e.g., ref [11]) or the adiabatic (see, e.g., ref [12]) implementation. The complete Hamiltonian including the non-adiabatic coupling between two degenerate electronic states then reads:

$$\hat{H} = \begin{pmatrix} T_{11} & K_{12} \\ -K_{12} & T_{22} \end{pmatrix} + \begin{pmatrix} V_{11} & 0 \\ 0 & V_{22} \end{pmatrix} \quad . \tag{11.21}$$

where $T_{11} = T_{22}$ denotes the kinetic part of the Hamiltonian and $V_{11,22}$ the adiabatic potentials. K_{12} is the non-adiabatic coupling (NAC) term formulated in Cartesian nuclear coordinates x_j

$$K_{12} = -\sum_j \frac{1}{m_j}\left(f_{12}^{(j)}\partial_{x_j} + \frac{1}{2}g_{12}^{(j)}\right) \tag{11.22}$$

with the mass m_j of the jth atom and using the notation ($\partial_x = \partial/\partial x$) for convenience. The first- and second-order derivative (non-adiabatic) coupling elements for the electronic wave functions $\Phi_{1,2}$ are given by

$$f_{12}^{(j)} = \langle \Phi_1 \mid \partial_{x_j} \Phi_2 \rangle \quad \text{and} \quad g_{12}^{(j)} = \langle \Phi_1 \mid \partial_{x_j}^2 \Phi_2 \rangle \ . \tag{11.23}$$

With K_{12} included, the complete nuclear Schrödinger equation is solved for the coupled states and all quantum effects like interferences or phase effects are included, provided one keeps track of the phases of the electronic wave functions. The NACs are calculated like the PES in the full-dimensional space and then projected onto the subspace spanned by the reactive coordinates [13]. This leads to a consistent description of the PES and NACs and includes contributions from full dimensionality. The resulting NACs $\tilde{f}_{12}^{r}$ and $\tilde{f}_{12}^{\varphi}$ (already including the masses) typically exhibit a spiky behaviour, requiring many grid points for the wave packet propagation and making the calculations quite time consuming.

Besides the spiky behavior, another much more fundamental issue has to be taken into account which is connected to the second derivative terms g_{12}^{j} of the non-adiabatic coupling. This problem has already been introduced and treated in [14, 15], but due to its importance it is briefly outlined again here. The second derivative terms g_{12}^{j} are mostly much smaller than the first derivative terms and hence are usually believed to be negligible. However, their omission will lead to a non-Hermitian Hamiltonian due to a non-Hermitian coupling

$$K^{(j)} = \tilde{f}_{12}^{(j)} \partial_{q_j} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} . \tag{11.24}$$

The hermiticity of $K^{(j)}$ requires

$$K_{12}^{(j)} = \tilde{f}_{12}^{(j)} \partial_{q_j} \stackrel{!}{=} -(K_{12}^{j})^{\dagger} . \tag{11.25}$$

The adjunct of an operator $A = BC$ with $B^{\dagger} = B$ and $C^{\dagger} = -C$ is

$$A^{\dagger} = -BC + [B, C], \tag{11.26}$$

and requiring A to be anti-Hermitian ($A = -A^{\dagger}$) is equivalent to the commutator equalling zero. In the actual case $B = \tilde{f}_{12}^{(j)}$ and $C = \partial_{q_j}$ do not commute, thus $K_{12}^{(j)}$ is not anti-Hermitian and therefore the total Hamiltonian is not Hermitian. To avoid this artifact, the non-Hermitian part can be compensated by anti-symmetrizing $A = \tilde{f}_{12}^{(j)} \partial_{q_j}$:

$$\begin{aligned} \tilde{A} &= \frac{A - A^{\dagger}}{2} = A - \frac{1}{2}[B, C] \\ &\Rightarrow \tilde{A}^{\dagger} = -\tilde{A} \ . \end{aligned} \tag{11.27}$$

In the present case, the commutator [B,C] results in:

$$\begin{aligned} [f, \partial_q]\Psi &= f\partial_q \Psi - (\partial_q f)\Psi - f\partial_q \Psi \\ &= -(\partial_q f)\Psi \end{aligned} \tag{11.28}$$

and we arrive at the extension

$$\tilde{f}_{12}^{(j)} \partial_{q_j} \rightarrow \tilde{f}_{12}^{(j)} \partial_{q_j} + \frac{1}{2} \partial_{q_j} \tilde{f}_{12}^{(j)} \tag{11.29}$$

leading to

$$\tilde{K}_{12} \approx \sum_j \left(\tilde{f}_{12}^{(j)} \partial_{q_j} + \frac{1}{2} \partial_{q_j} \tilde{f}_{12}^{(j)} \right) \ . \tag{11.30}$$

Actually (11.29) is the first part of the decomposition of the second derivative term

$$g_{12}^{(j)} = \partial_{x_j} f_{12}^{(j)} + h_{12}^{(j)}; \tag{11.31}$$

thus Eq. (11.30) only neglects the Hermitian part $h_{12}^{(j)} = \langle \partial_{x_j} \Phi_1 | \partial_{x_j} \Phi_2 \rangle$ of $g_{12}^{(j)}$ and provides a very efficient way to calculate Hermitian non-adiabatic couplings.

11.3 Choice of Dynamical Coordinates and Hamiltonian in Reduced Dimensionality

To minimize the loss of information that comes with the reduction of dimensionality, process-adapted coordinates are needed, which describe the relevant structural changes and contain all modes that are active on the intrinsic timescale of the reaction. This approach is particularly suited for the femtosecond time regime. Typically only few fast reactive motions dominate the molecular process here. In the following, we will describe three different approaches developed in our group, ranging from manual to automated selection of the reduced reactive coordinates.

11.3.1 Manual Selection by Chemical Intuition

During a chemical reaction, the molecular geometry continuously changes, making the often used normal mode expansion unattractive as, typically, a large number of normal modes is needed. Instead, the use of process adapted (reactive) coordinates is more promising. Note that each reactive coordinate includes projections on several normal modes. As a well-studied text-book example, we discuss the photoinduced electrocyclic ring-opening in cyclohexadiene (CHD) [13–15]. The manually performed selection procedure was based on a thorough analysis of the geometrical structures, which were known from quantum chemistry to be important in the full coordinate space. These characteristic structures are the ground state geometry of cyclohexadiene and hexatriene as well as the geometries of the conical intersections that mediate the ultrafast transition from the S_1 state to the ground state [16–19]. Including all important structures, an appropriate coordinate space was constructed using the modes depicted in Figure 11.1 (left).

Three C_2-symmetry conserving modes, describing the structural changes from CHD to hexatriene, are involved: The torsions β and γ as well as the bond length R of the σ-bond where the ring opening occurs. The remaining asymmetric motion is represented by the angle α and needs to describe the structural changes along the relevant conical intersection seam. For the dynamical

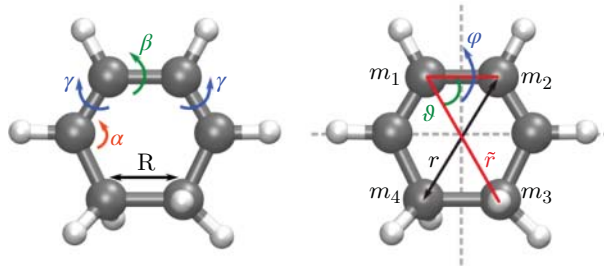


Figure 11.1 Reduced set of coordinates for the ring opening of CHD with a detailed explanation given in the text.

description of the system, these coordinates are not suitable due to their high degree of kinetic coupling. Thus, they are transformed into a set of only three coordinates (see Figure 11.1, right) reducing, at the same time, the complexity in the kinetic Hamiltonian. The asymmetric coordinate is now represented by the length of the diagonal r using a linear approximation for the angle α whereas the second diagonal $\tilde{r}$ is kept constant. The C_2-symmetric angle ϑ describes a scissoring motion leading to a stretching of the σ-bond. Finally, the remaining symmetric coordinate φ distorts the molecule further away from planarity comprising both the torsions β and γ.

The corresponding kinetic Hamiltonian is derived from the full Cartesian nuclear kinetic energy operator:

$$\hat{T}_x = -\frac{\hbar^2}{2}\sum_{i=1}^{3N}\frac{1}{m_i}\frac{\partial^2}{\partial x^2}. \tag{11.32}$$

First, all H-atoms are assigned to the corresponding C-atoms and the masses of the remaining six atoms are contracted to four points of mass according to the broken lines in Figure 11.1, right. Thus, the masses of the C-atoms located on the horizontal dashed line are split between the two adjacent atoms. Eliminating the center of mass motion and transforming to mass-weighted Cartesian Jacobi coordinates leaves nine degrees of freedom, which are then transformed to the previously selected reactive coordinates. These considerations finally yield a reduced Hamiltonian in three dimensions

$$\hat{T} = -\frac{1}{2}\left\{\frac{B}{r}\partial_r^2 r + \left(\frac{B}{r^2}+\frac{\tilde{B}}{\tilde{r}^2}\right)\left(\frac{1}{sin^2\vartheta}\partial_\varphi^2 + \partial_\vartheta^2 + \cot\vartheta\,\partial_\vartheta\right)\right\}. \tag{11.33}$$

The kinetic energy operator $\hat{T}$ resembles a combination of two operators in spherical harmonics, sharing the same angular variables ϑ and φ. As $\tilde{r}$ is kept constant, its derivative terms vanish. The inverse reduced masses B and $\tilde{B}$ are given by

$$B = \frac{1}{m_2}+\frac{1}{m_4}, \quad \tilde{B} = \frac{1}{m_1}+\frac{1}{m_3} \tag{11.34}$$

with $m_1 = m_2 = 1.5(m_C + m_H)$ and $m_3 = m_4 = 1.5m_C + 2.5m_H$. For details of the derivation see [13]. For the example of the electrocyclic ring-opening of CHD, the resulting PES for ground and first excited state include the equilibrium geometries of reactant (cyclohexadiene) and product (cZc-hexatriene) states, an extended Franck–Condon (FC) region, and energetically accessible conical intersection [13, 15] between excited and ground state.

The quantum dynamical calculations performed for the photoinduced electrocyclic ring-opening of CHD in the adiabatic representation in reduced dimensionality were well able to reproduce the observed overall timescale as well as the product distribution [14, 15]. This good agreement of the computational results with the experimental data validated the concept of reactive coordinates and their manual selection as outlined above.

11.3.2 The *G*-Matrix Formalism

The transformation of the kinetic operator from Cartesian to spherical coordinates was not straightforward and is quite specific to the given molecular process. Since we would like to employ reactive coordinates, which are very specific to the chemical process in question, it is desirable to develop a general formalism which allows for a convenient transformation of the kinetic energy operator to any arbitrary set of coordinates. The *G*-matrix formalism [20, 21] introduced in this section provides a way to achieve this goal.

11.3.2.1 General Setup

Originally introduced by Podolsky [22] in 1928, the fundamental step is a Jacobian transformation from Cartesian coordinates x_i to a new set q_r via

$$q_r = \sum_{i=1}^{3N} j_{ir}^{-1} x_i. \tag{11.35}$$

It uses the Jacobian Matrix

$$j_{ir} = \frac{\partial x_i}{\partial q_r} \tag{11.36}$$

containing all partial first derivatives. The metric tensor g_{rs} of the new set of coordinates is thus given by

$$g_{rs} = \sum_{i=1}^{3N} j_{ri} j_{is} = \sum_{i=1}^{3N} \frac{\partial x_i}{\partial q_r} \frac{\partial x_i}{\partial q_s}. \tag{11.37}$$

Denoting g as the determinant of g_{rs}, the kinetic operator $\hat{T}_q$ can be written as [22]

$$\hat{T}_q = -\frac{\hbar^2}{2\mu} \sum_{r=1}^{M} \sum_{s=1}^{M} g^{-1/4} \frac{\partial}{\partial q_r} \left(g^{1/2} g^{rs} \frac{\partial}{\partial q_s} g^{-1/4} \right). \tag{11.38}$$

Here, g^{rs} is the cofactor matrix of g_{rs}, and the mass μ is identical for all particles. Obviously, it is desirable to be able to attribute different masses for different particles. Wilson and co-workers [20] thus write the G-matrix as

$$G_{rs}(q) = \sum_{i=1}^{3N} \frac{1}{m_i} \frac{\partial q_r}{\partial x_i} \frac{\partial q_s}{\partial x_i}, \tag{11.39}$$

where the summation is conducted over all $3N$ degrees of freedom of each atom i with its mass m_i, and the G-matrix G_{rs} is symmetric ($G_{rs} = G_{sr}$).

The terms required to set up $\hat{T}_q$ within Eq. (11.38), and thus the G-matrix in Eq. (11.39), are dependent on the coordinates q. Assuming that the Jacobian determinant is constant [21] and the coordinate dependence can be disregarded, Eq. (11.38) simplifies to

$$\hat{T}_q \simeq -\frac{\hbar^2}{2} \sum_{r=1}^{M} \sum_{s=1}^{M} \frac{\partial}{\partial q_r} \left[G_{rs} \frac{\partial}{\partial q_s} \right]. \tag{11.40}$$

We therefore arrive at a very general formulation of the kinetic energy operator $\hat{T}_q$, which is expressed in arbitrary coordinates q and can be used after a direct transformation from Cartesian coordinates. In contrast to $\hat{T}_x$ (Eq. (11.32)), $\hat{T}_q$ contains cross partial derivatives with respect to q_r and q_s, which represent a kinetic coupling between two coordinates. These cross terms originate from non-orthogonal connections of Cartesian components and mean that a motion along q_r induces a motion along q_s with a certain amplitude corresponding to the degree of the kinetic coupling. If we consider linear reactive coordinates, the diagonal elements of G_{rs} can be regarded as the reciprocal reduced mass along this motion, equivalent to the reduced mass of molecular normal modes [20]. It is also straightforward to set up G_{rs} for non-linear coordinates (i.e., the molecular motion changes between different regions of the coordinate). In these cases, the elements of G_{rs} are no longer constant along a given coordinate but can be set up in the same way as for linear coordinates. In some cases, the use of non-linear coordinates violates the approximation of the Jacobian determinant j being constant, which can lead to artifacts in the wave packet simulations. As outlined in ref. [23], this problem can be handled by rescaling the reactive coordinates q such that the change in volume due to coordinate transformation – and thus j – is constant at all points along the coordinate.

11.3.2.2 Practical Computation of the *G*-Matrix Elements

In practice, the *G*-matrix is more accessible and easier to calculate via its inverse elements [24]:

$$G_{rs}^{-1} = \sum_{i=1}^{3N} m_i \frac{\partial x_i}{\partial q_r} \frac{\partial x_i}{\partial q_s}. \tag{11.41}$$

The dependence of variables is reversed, and the elements of G_{rs}^{-1} can simply be calculated via finite differences by representing the molecule in the Cartesian space and displacing it along the reactive coordinates q_r and q_s. If the *G*-matrix is chosen to be accessed via its inverse elements, the coordinate transformation must be unique – i.e., it must be reversible – for the *G*-matrix to be invertible. One crucial criterion for this is that the Eckart conditions [25] are fulfilled, which ensure the separation of internal from external coordinates. This means that with the displacement of atoms along the reactive coordinates, the center of mass must not change, and the molecule must not gain angular momentum. The first Eckart condition excludes translational motion by requesting that the mass weighted scalar products of atomic positions $x_i(q)$ and displacement vectors $x_i(q + dq)$ cancel each other out:

$$\sum_{i=1}^{N} m_i [x_i(q) \cdot x_i(q + dq)] = 0. \tag{11.42}$$

This can be ensured by putting the molecular center of mass in the origin at all points along the coordinate. The second Eckart condition

$$\sum_{i=1}^{N} m_i x_i(q) \times x_i(q + dq) = 0 \tag{11.43}$$

excludes rotational degrees of freedom and can be fulfilled by applying a pseudo-rotation procedure as outlined in ref. [26].

11.3.2.3 Photorelaxation of Uracil in Linear Reactive Coordinates

In this section, a practical example of using reduced linear reactive coordinates along with the *G*-matrix is given. The intention was to model the photophysical relaxation of the RNA nucleobase uracil out of the excited state S_2 and steer it by light [27]. By performing wave packet simulations on PESs, the photochemical pathway of the molecule, as well as excited state lifetimes, can be extracted. From various stationary quantum chemistry calculations of the excited state profile, three important molecular structures were identified to play a major role during the relaxation process: the FC point where the wave packet enters the excited state surface, a conical intersection where the wave packet decays to the lower-lying S_1 state, and a local intermediary energetic minimum in the S_2 state. As three non-collinear points in space span a plane, these three structures can be used to construct a two-dimensional coordinate space. Specifically, the displacement vector from the FC point geometry x_{FC} to the optimized conical intersection x_{CoIn} was taken as the first coordinate $q_1 = x_{\mathrm{CoIn}} - x_{\mathrm{FC}}$. The second coordinate $q_2 = x_{\mathrm{S_2min}} - x_{\mathrm{FC}}$ was obtained through the displacement vector from x_{FC} to the local S_2 minimum $x_{\mathrm{S_2min}}$ and subsequent orthogonalization to q_1. Both q_1 and q_2 are shown in Figure 11.2.

Using quantum chemical calculations on the MRCI/CASSCF level of theory, energy values for ground and excited states can be assigned to every molecular structure spanned by the combined displacement along q_1 and q_2, yielding PES for the adiabatic electronic states. The PES for the ground state and the second excited state (bright state) of uracil are also illustrated in Figure 11.2. Along with the energy, transition dipole moments that are necessary for laser excitation and the NACs between S_2 and S_1 (cf. section 11.2.4) were also obtained from quantum chemistry. All these

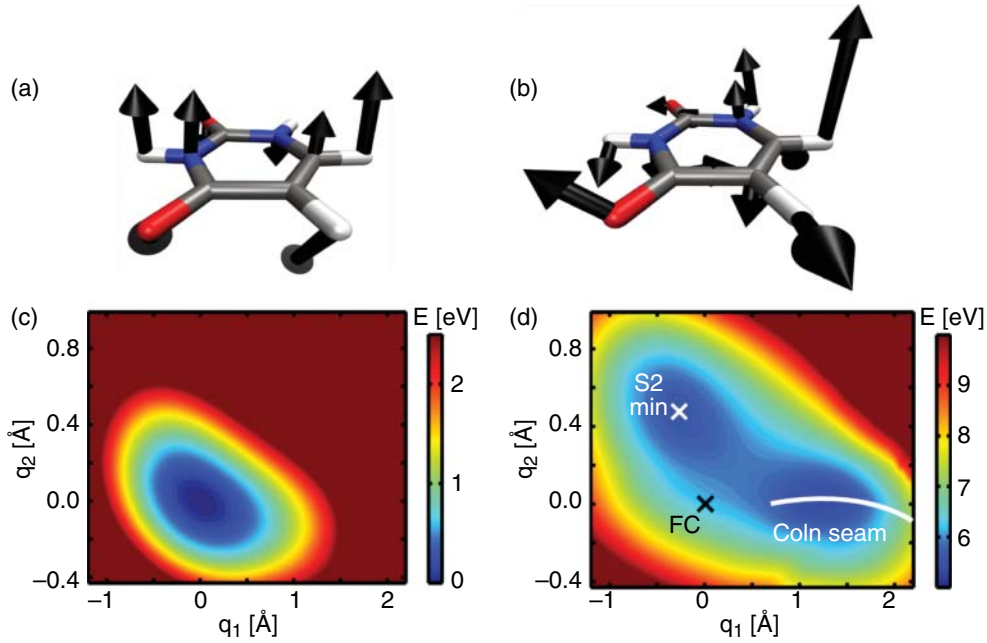


Figure 11.2 Two-dimensional representation of uracil. The displacement vectors q_1 and q_2 are shown in (a) and (b) respectively. The PES of the S_0 and S_2 electronic states according to MRCI/CASSCF calculations performed for combined displacements of q_1 and q_2 are shown in (c) and (d). The three specified molecular geometries in (d) were used to construct the displacement vectors to set up the low-dimensional representation.

elements are part of the potential energy operator $\hat{V}$ in the TDSE (11.3), and the setup of the kinetic energy operator $\hat{T}_q$ in reduced coordinates was conducted with the *G*-matrix procedure described in the previous section. As q_1 and q_2 are linear reactive coordinates, the *G*-matrix elements are constant along both dimensions, and the *G*-matrix itself used in Eq. (11.40) and noted in atomic units takes the following form:

$$G_{q_1q_2} = \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix} = 10^{-4} \cdot \begin{pmatrix} 1.67 & -0.70 \\ -0.70 & 1.39 \end{pmatrix}. \tag{11.44}$$

As described above, $G_{q_1q_2}$ was obtained by calculating its inverse matrix according to Eq. (11.41) by finite differences after displacing the molecular structure represented in Cartesian space along both coordinates. The diagonal elements of $G_{q_1q_2}$ can be regarded as the reciprocal reduced mass along the respective coordinate, which is 3.28 atomic mass units for a vibration along q_1 and 3.95 atomic mass units for a vibration along q_2 respectively. The associated frequencies are 900 cm^{-1} for q_1 and 480 cm^{-1} for q_2, which is similar to common angular or out-of-plane vibrations found in small organic molecules. The off-diagonal element q_{12} is the value of the kinetic coupling between q_1 and q_2, and $G_{q_1q_2}$ is symmetric ($q_{12} = q_{21}$). Using this simulation setup, the performed wave packet simulations were able to correctly reproduce the experimentally measured excited state lifetimes [27, 28]. The setup of the *G*-matrix demonstrated here is straightforward and can be translated to any other reasonable linear coordinate space or molecule. For non-linear coordinates, the only difference is that the matrix elements are not constant along the specific dimensions and must be computed at each grid point, according to the local displacement.

11.3.3 Automatic Generation of Linear Coordinates

We now introduce two methods to automate the search for linear coordinates as described in ref [29]. These techniques are also appropriate for finding low-dimensional coordinate spaces in cases where chemical intuition reaches its limits. This can be, for example, a ground state reaction taking place on slower timescales (picosecond and longer) and involving many normal modes.

At the start of all further considerations, the critical points of the reaction, such as reactant, product and transition state need to be identified. The vectors connecting the molecular structures at these points already span a reduced-dimensional space of linear coordinates that will be extended with two different techniques in the following. The first one uses the full-dimensional intrinsic reaction coordinate (IRC) [30] to map the reaction path. The IRC is defined as a path of minimum energy on the PES that connects the reactant geometry with any transition states and the product. The second method presented in this section employs semi-classical trajectories to sample the configuration space of the reaction. Both techniques are summarized as flowcharts in Figure 11.3.

11.3.3.1 IRC Based Approach

The IRC based approach presents a relatively simple method to automate the search for linear coordinates. First, the IRC of the reaction is calculated using conventional quantum-chemical methods.

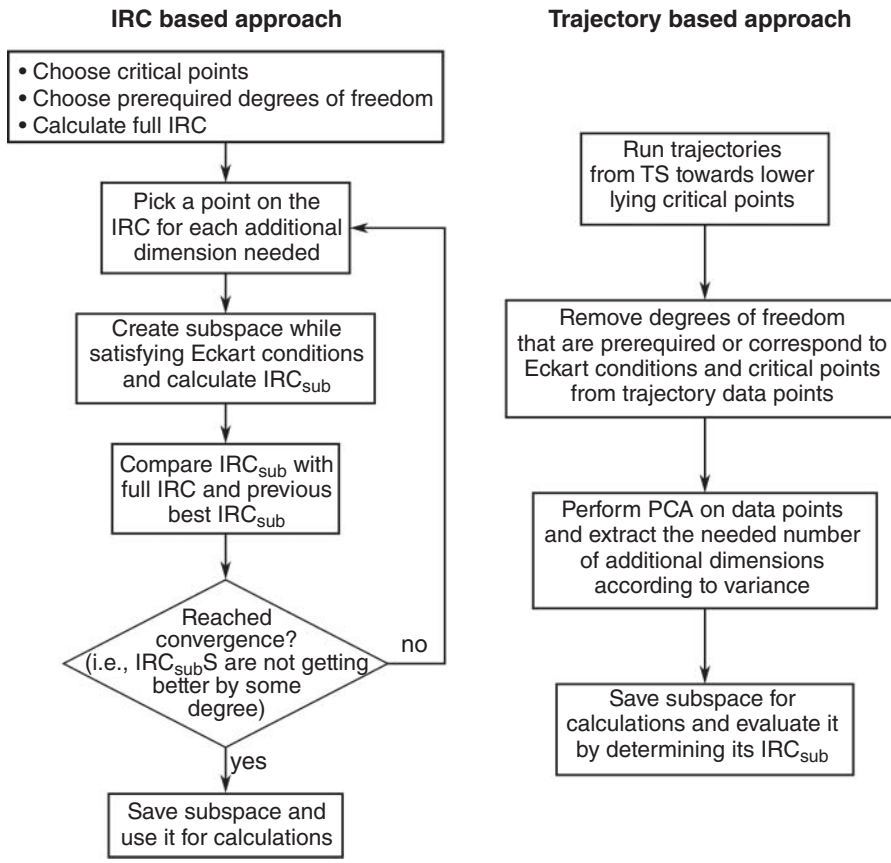


Figure 11.3 Flowcharts illustrating the IRC based and trajectory-based approaches for the semi-automatic search for reduced linear coordinate spaces.

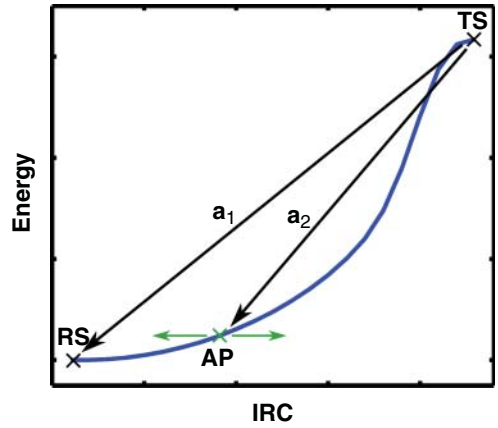


Figure 11.4 One possibility to construct IRC based subspaces. Additional to the transition state (TS) and reactant state (RS), an additional point (AP) is chosen along the IRC. The three corresponding geometries are used to construct to the vectors $\vec{a}_i$ pointing from the TS to the other two structures and spanning the new subspace.

The goal of the dimensionality reduction is to reproduce this non-linear IRC as well as possible within the new set of linear coordinates. Additional coordinates to the ones determined by the critical points are defined by segmenting the IRC with additional points as shown in Figure 11.4.

The number of possible reduced-dimensional subspaces N_{ss} is technically only limited by the number of IRC points N_{IRC} and dimensions D:

$$N_{ss} = \frac{N_{IRC}!}{(N_{IRC} - D)!}. \tag{11.45}$$

However, the number of actual subspaces can be reduced by omitting points that are close to the critical points and, in the case of more than one additional dimension, to each other. A suitable strategy to further reduce the number of subspaces is to start with a coarse segmentation of the IRC, identify the best subspaces and optimize them by finer segmentation near the corresponding points.

The basis vectors for reduced-dimensional subspaces are constructed by connecting a reference structure with the predetermined critical points of the reaction and the additional points along the IRC. In this example, the transition state structure is used as a reference (Figure 11.4), but there are of course other ways to construct the basis vectors by choosing different reference structures. Once the basis vectors have been constructed, any external degrees of freedom need to be removed from the subspaces by satisfying the Eckart conditions [25] for all molecular structures with respect to the reference structure.

Finally, the constructed subspaces need to be tested for their ability to reproduce the reaction pathway. This can be achieved by calculating a low-dimensional IRC_{sub} within the newly constructed subspaces and comparing it to the full-dimensional IRC. Note that IRC_{sub} is *not* the projection of the full-dimensional IRC onto the subspace as illustrated in Figure 11.5.

The IRC_{sub} can be compared to the full-dimensional IRC by using a distance criterion ΔA^i as well as an energy criterion ΔB^i as quality measures for the ith subspace. For the former, the distance $\Delta r(IRC^i_{sub})$ from each point of the IRC_{sub} to its full-dimensional counterpart is evaluated. Integration over all points yields the distance quality criterion ΔA^i:

$$\Delta A^i = \int \Delta r(IRC^i_{sub}) \mathrm{d}IRC^i_{sub}. \tag{11.46}$$

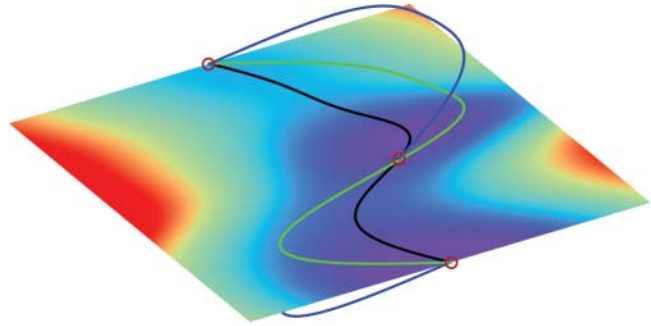


Figure 11.5 Two-dimensional affine subspace of a three-dimensional potential spanned by three critical points (red circles). The color map of the two-dimensional slice encodes the potential energy landscape ranging from high energy (red) to low energy (purple) regions. It is overlaid with the full-dimensional IRC (blue), its projection onto the subspace (green) and the reduced-dimensional IRC_{sub} (black).

The second quality measure is calculated analogously but uses the energy difference $\Delta E(\mathrm{IRC}^i_{\mathrm{sub}})$ between each point of the $\mathrm{IRC}_{\mathrm{sub}}$ and the full-dimensional IRC:

$$\Delta B^i = \int \Delta E(\mathrm{IRC}^i_{\mathrm{sub}})\mathrm{dIRC}^i_{\mathrm{sub}}. \tag{11.47}$$

Finally, a combined quality criterion Q^i can be formulated by averaging over both quantities:

$$Q^i = \frac{N\Delta A^i}{\sum_{j=1}^{N}\Delta A^j} + \frac{N\Delta B^i}{\sum_{j=1}^{N}\Delta B^j}. \tag{11.48}$$

In this way, the subspaces can be directly compared, where a smaller Q^i means better reproduction of the full-dimensional IRC and, therefore, a better subspace i.

11.3.3.2 Trajectory-Based Approach

The second approach to construct linear coordinates employs semi-classical trajectories to sample the configuration space of the reaction. A good starting point is to calculate trajectories from the area around the transition state towards the product and extract geometries at regular intervals to obtain a data set that maps the reaction space. This set of molecular structures can be subsequently decomposed using principal component analysis (PCA) to extract the most important degrees of freedom and thus a set of linear coordinates for quantum dynamics.

The basic idea of PCA is to find a set of orthogonal dimensions and sort them by the variance of data along them. For example, in Figure 11.6 the variance of the blue dataset is highest along $\vec{u}$ and second highest along $\vec{v}$. The vectors $\vec{u}$ and $\vec{v}$ are called the principal components of the dataset.

For higher-dimensional data, it is often sufficient to project the data points onto the first two or three principal components and discard the others, where the variance becomes negligibly small. PCA is a widely used procedure in data science for dimensionality reduction, feature extraction and visualization of large data sets. For an arbitrary data set $\mathbf{X}$ with M data points, it works by computing the covariance matrix $\mathbf{C}$:

$$\mathbf{C} = \frac{1}{M-1}(\mathbf{X} - \bar{\mathbf{X}})(\mathbf{X} - \bar{\mathbf{X}})^T. \tag{11.49}$$

$\mathbf{C}$ contains the variance of the data along dimension i as diagonal elements x_{ii} and the covariance between dimensions i and j as off-diagonal elements x_{ij}. PCA now aims to eliminate correlation between the dimensions, so that the off-diagonal elements become zero and $\mathbf{C}$ is diagonal. The

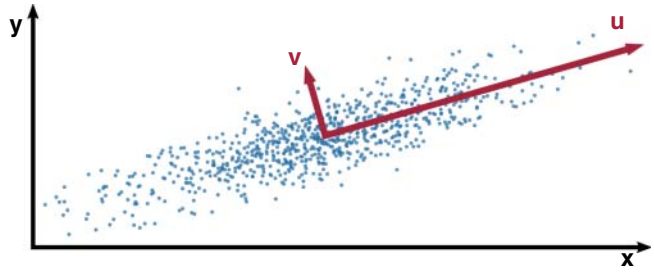


Figure 11.6 Principal component analysis of the blue dataset. The vectors $\vec{u}$ and $\vec{v}$ are the principal components of the two-dimensional distribution and scaled by the variance of the data along them.

eigenvectors of **C** obtained in the process of diagonalization are the principal components of the data and the corresponding eigenvalues give the variance along these new dimensions.

Using PCA on trajectory data is an established technique [31–34], its application to construct reduced-dimensional PES for quantum dynamical simulations is, however, quite recent [29]. In this context, a raw data matrix **Y** contains the geometries of an n-atomic molecule along the trajectories as row vectors in the format $(x_1, y_1, z_1, x_2, y_2, z_2, \ldots, x_n, y_n, z_n)$. It is important that all data points satisfy the Eckart conditions [25] with respect to a given reference point (e.g., the transition state of the reaction), to remove any rotational or translational degrees of freedom that would skew the results of the PCA. The data is subsequently centered around the Cartesian origin by shifting each column vector by the value of its mean, yielding a matrix $\tilde{\mathbf{Y}}$. The subspace containing the critical points is removed by constructing a matrix **V** that contains the orthonormal basis vectors of the subspace as row vectors and calculating the final data matrix **X**:

$$\mathbf{X} = \tilde{\mathbf{Y}} - (\tilde{\mathbf{Y}} \cdot \mathbf{V}^T)\mathbf{V}. \tag{11.50}$$

The PCA is then applied to **X** and the principal component vectors with the highest eigenvalues along with the row vectors of **V** can be used as basis vectors $\vec{q}_i$ for the construction of a reduced-dimensional coordinate space. The quality of this subspace is evaluated according to Eqs. (11.46) to (11.48), analogously to the IRC-based method.

11.3.3.3 Comparison of Both Techniques for Linear Subspaces

As an example, to compare the linear coordinate spaces obtained with both techniques, the proton transfer between the oxygen atoms of (*Z*)-hydroxyacryloyl chloride will be discussed [35–37]. Modelling a reaction such as this is a typical task in quantum dynamics. The reaction takes place in the ground state, is therefore slower, and involves more normal modes than an ultrafast photochemical process. Both the IRC and the trajectory based method were used to reduce the number of dimensions from 27 to 3 [29]. The IRC of the reaction is shown along with the structures of the critical points in Figure 11.7(a).

For the IRC based approach, the part of the IRC between reactant and transition state was segmented with 245 points with a distance of 0.0028 Å to each other. Starting with the sixth point, a subspace was created every ten IRC points using the transition state structure as a reference. The quality of the resulting subspaces was evaluated with the criteria defined in Eqs. (11.46) to (11.48).

For the trajectory-based approach, 50 semi-classical trajectories were calculated. Starting structures were generated by displacing the transition state geometry in steps of 0.002 Å along the imaginary mode towards the reactant state. Each trajectory was propagated without initial momentum for a total of 100 fs, during which all trajectories passed the reactant structure. Subsequently, the PCA was employed to obtain a subspace that was evaluated according to the same

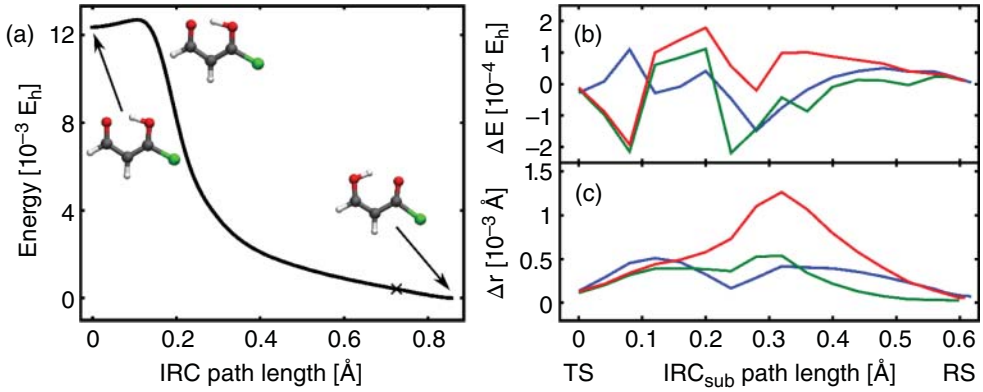


Figure 11.7 Energy profile (a) along the IRC of the proton transfer reaction in (Z)-hydroxyacryloyl chloride. The product is shown on the left, the transition state in the middle and the reactant on the right. The cross marks the IRC point which yielded the best subspace in the IRC based method. Reduced-dimensional subspaces aim to represent the part between transition state and reactant. Energy difference $\Delta E(IRC^i_{sub})$ and distance functions $\Delta r(IRC^i_{sub})$ between the IRC of the subspace and the full-dimensional IRC are plotted in (b) and (c) respectively to provide a quality measure for the subspaces. The red curve corresponds to the trajectory based method ($\Delta Q^i = 2.438$), while the green ($\Delta Q^i = 1.233$) and blue ($\Delta Q^i = 0.510$) curves belong to selected subspaces obtained with the IRC based method.

quality criteria as the ones generated with the IRC based method. Selected results are compiled in Figure 11.7. For a more extensive comparison, please refer to the original publication [29].

Figure 11.7 shows that the energy difference is mostly dominated by noise in both methods, discernible by the fluctuations around 0 E_h. This can be explained by the fact that the proton transfer is well described in three dimensions and the IRC_{sub} follow the full-dimensional IRC very closely. The main reasons for energy deviations in this case are discretization errors that can be reduced by employing a finer grid for the IRCs.

Another observation is that the distance function is larger for the trajectory-based method than for the best subspaces obtained with the IRC based technique. Because the latter approach employs an additional point on the IRC, the resulting IRC_{sub} is bound to be closer to the full-dimensional IRC than one calculated with the trajectory-based method.

Finally, the coordinate vectors spanning the three-dimensional subspaces obtained with both methods are visualized in Figure 11.8. The two vectors on the left show the displacement from the

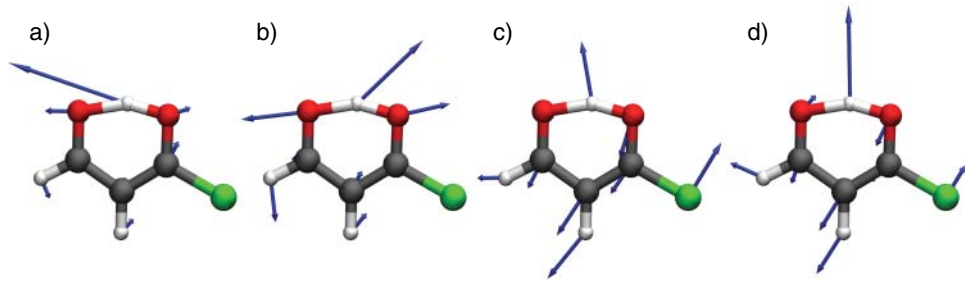


Figure 11.8 Basis vectors spanning three-dimensional subspaces for the proton transfer reaction in (Z)-hydroxyacryloyl chloride. On the left the displacement vectors from the transition state to the reactant (a) and the product (b) are shown. The third coordinate was determined with the trajectory-based approach (c) and the IRC- based method (d) respectively.

transition state to the reactant (a) and the product (b), respectively. These two basic coordinates are the same in both methods. On the right, the best third coordinate obtained with the IRC based method (c) and the one obtained with the trajectory-based method (d) are shown. Both approaches yield a very similar third coordinate and are, therefore, very comparable in accuracy. In both cases, this third coordinate contains a rather unintuitive motion of atoms not directly involved in the reaction as well as an in-plane movement of the proton away from the ring. As such a coordinate would probably be omitted in manual selection procedures, this example illustrates the importance of unbiased, automated techniques to construct coordinate spaces for quantum dynamics.

11.3.4 Automatic Generation of Non-Linear Coordinates

In grid-bases quantum dynamics, the distance between two grid points corresponds to a Cartesian displacement vector between two molecular geometries. This displacement vector will be the same across the grid in a linear subspace, whereas a non-linear subspace means that the displacement vector can change at every grid point. Quantum dynamics in non-linear subspaces typically requires fewer coordinates than in linear subspaces to describe the same reaction. The technique presented in this section provides a way of constructing generalized non-linear subspaces. For this, we will employ an artificial neural network, specifically (but not exclusively) designed for dimensionality reduction—an autoencoder [38].

Autoencoders consist of two parts (Figure 11.9): an encoder that compresses the input into a central, reduced-dimensional representation, called the code layer, and a decoder that attempts to reconstruct the input from the information in the code layer. The bottleneck architecture, sometimes referred to as an undercomplete autoencoder, drives the dimensionality reduction, as the number of neurons in the code layer determines the maximum number of features to learn. By minimizing the error between the reconstruction and the original input, the neural network should learn the best reduced-dimensional representation of the input.

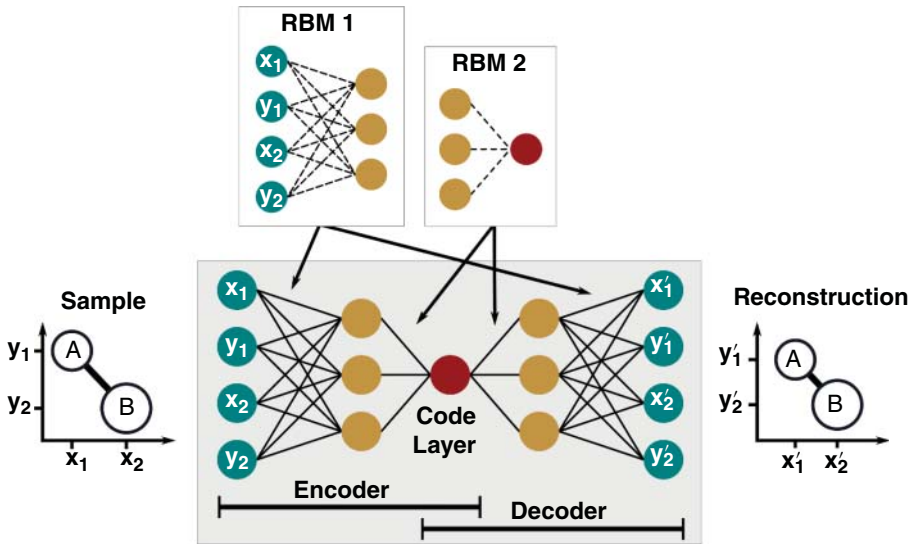


Figure 11.9 Schematic representation of an autoencoder. Each Cartesian coordinate of a diatomic molecule AB constitutes a green input node on the left. The network learns a one-dimensional representation (red) of the original data on the left and uses it to reconstruct the input. Restricted Boltzmann machines (RBMs) are used to pre-train the weights and biases for the actual autoencoder.

The learning process itself works by optimizing the connections between the various neurons. In this context, the activation z_j of a neuron j in the current layer of the neural network is determined by the weighted sum of the activation values x_i from the previous layer:

$$z_j = \sigma_j \left(\sum_{i=1}^{n} w_{ij} x_i + b_j \right). \tag{11.51}$$

The weights w_{ij} define how strongly two neurons i and j interact, i.e., how much the activation of i contributes to that of j. Additionally, a bias b_j shifts the activation value of the neuron and provides a threshold for meaningful activation. For instance, a bias of −5 means the neuron only activates when the weighted sum of activations from the previous layer is greater than 5. The resulting value is mapped onto a predefined data range with an activation function $\sigma_j(u)$ that also introduces the non-linearity between the layers. In the example discussed below, σ_j will be a sigmoid function

$$\sigma_j(u) = \frac{1}{1 + \exp(u)} \tag{11.52}$$

that maps the activation onto the range between 0 and 1 for all hidden layers. During training, the weights and biases of the autoencoder are optimized to minimize the loss function L. In this example, L is the mean squared error between the input coordinates $\vec{x}_i^{\,\mathrm{inp}}$ and the reconstructed output $\vec{z}_i^{\,\mathrm{rec}}$ for N training samples:

$$L = \frac{1}{N} \sum_{i=1}^{N} (\vec{x}_i^{\,\mathrm{inp}} - \vec{z}_i^{\,\mathrm{rec}})^2. \tag{11.53}$$

Using Eqs. (11.51) and (11.52), the chain rule is applied to Eq. (11.53) to calculate the gradient of L with respect to the weights w_{ij} and biases b_j. This procedure is known as backpropagation [39] and has become one of the most popular algorithms to train artificial neural networks.

In practice, especially for autoencoders with many hidden layers, training can be slow or even fail to converge if the initial guess for weights and biases differs strongly from a useful solution. This problem can be handled by employing a set of smaller neural networks, Restricted Boltzmann machines (RBMs), to pre-train the autoencoder [40]. An RBM only contains an input layer and one hidden layer and learns a probability distribution over its input nodes by optimizing the weights and biases between the two layers. The initial weights for an RBM can be chosen as a normal distribution and the biases can simply be initialized as zeros. Once an RBM is trained, the activations of its hidden layer can be used as input in the next RBM until all RBMs are trained. The number of neurons per layer in each consecutive RBM is chosen to reflect the architecture of the autoencoder (see also Figure 11.9), so that the weights and biases of the trained RBMs can enter as initial guesses in the actual training routine of the autoencoder.

However, before the training can begin, the autoencoder needs a dataset whose features it can learn. In this case, the data is composed of molecular geometries that sample the reaction space. Such structures can be generated by calculating the IRC of the reaction and exploring the space around it via a swarm of semi-classical trajectories from regularly spaced points along the IRC. For this approach to work, the trajectories need to be constrained by an algorithm like RATTLE [41] to run orthogonal to the IRC and to prevent them from simply following the gradient descent. Furthermore, the Eckart conditions [25] need to be satisfied in every time step with respect to the starting structure of the respective trajectory to remove any translations and rotations that might have been introduced by the constraints.

The resulting set of molecular geometries typically contains a high degree of redundancy, especially when many trajectories are launched from the same starting point. This can be problematic,

as the unnecessarily large dataset slows down the training routine without adding information and might even skew the results if the redundancies are not evenly distributed. It is therefore advisable to clean the data by calculating the Euclidean distance between all data points and removing points that fall below a certain threshold. Finally, all data points should be normalized so that the components of every Cartesian coordinate vector are between 0 and 1, which generally improves training efficiency.

Once the autoencoder has learned a low-dimensional representation of this molecular dataset, it is time to generate a grid on which to perform quantum dynamics. The construction of a rectangular grid in a non-linear coordinate space can however be quite challenging, because the autoencoder coordinates can be highly curvilinear. It is easier to choose a linear subspace similar to that found by the autoencoder, construct a rectangular grid there and project it onto the non-linear subspace. For this purpose, a PCA can be performed on the training data to extract its principal components and use them as basis vectors for a rectangular grid. The encoder part of the trained autoencoder can now be employed to map the grid points onto the low-dimensional coordinate space given by the code layer and refine this projection using the decoder part. In the following, let $\vec{y}' = \mathbf{F}(\vec{y})$ be a mapping of the full-dimensional vector $\vec{y}$ onto a low-dimensional subspace. Its inverse mapping $\vec{y}'' = \mathbf{H}(\vec{y}\prime)$ transforms the reduced-dimensional vector $\vec{y}'$ to a full-dimensional vector $\vec{y}''$. Both mappings are ideal in the sense that the reconstruction $\vec{y}''$ equals the original vector $\vec{y}$ if $\vec{y}$ lies within the subspace and that the distance $\|\vec{y}'' - \vec{y}\|$ is as small as possible if $\vec{y}$ lies outside of the subspace. The encoder $\tilde{\mathbf{F}}$ and decoder $\tilde{\mathbf{H}}$ of the autoencoder are approximations to these optimal mappings $\mathbf{F}$ and $\mathbf{H}$. The trained autoencoder therefore produces the approximate projection $\vec{y}'' = \tilde{\mathbf{H}}(\tilde{\mathbf{F}}(\vec{y}))$. This allows the construction of a set of basis vectors $\vec{a}_j$ by adding a small perturbation $\varepsilon\vec{e}_j$ to the activations of each node j in the central code layer, where $\vec{e}_j$ is a unit vector and ε determines the magnitude of the perturbation:

$$\vec{a}_j = \tilde{\mathbf{H}}(\vec{y}' + \varepsilon\vec{e}_j) - \tilde{\mathbf{H}}(\vec{y}'). \tag{11.54}$$

This projection is approximate and can be refined by minimizing the distance $\|\tilde{\mathbf{H}}(\vec{y}') - \vec{y}\|$ with respect to variations in $\vec{y}'$. Therefore, the distance vector is projected onto the basis vectors $\vec{a}_j$ using the non-orthonormal projection matrix $\mathbf{A} = [\vec{a}_1, \vec{a}_2, \dots, \vec{a}_S]$ with S dimensions in the subspace. The first step of this iterative optimization is

$$\vec{y}'_1 = \vec{y}' + h\varepsilon(\mathbf{A}^T\mathbf{A})^{-1}(\vec{y} - \tilde{\mathbf{H}}(\vec{y}'))^T\mathbf{A}, \tag{11.55}$$

where h is the step size. The next steps are

$$\vec{y}'_k = \vec{y}'_{k-1} + h\varepsilon(\mathbf{A}^T\mathbf{A})^{-1}(\vec{y} - \tilde{\mathbf{H}}(\vec{y}'_{k-1}))^T\mathbf{A} \tag{11.56}$$

until $\|\tilde{\mathbf{H}}(\vec{y}') - \vec{y}\|$ is minimal and the rectangular PCA grid is completely mapped onto the non-linear subspace found by the autoencoder. Finally, a PES can be calculated on the non-linear grid and used in quantum dynamical simulations. The kinetic energy is represented with the G-matrix formalism described in section 11.3.2.

To evaluate the quality of this method, once again the proton transfer in (Z)-hydroxyacryloyl chloride will be used as a test system with the projection of the IRC onto the subspace to the full-dimensional IRC as a quality measure. Figure 11.10 illustrates the distance Δr of the projected IRC from its full-dimensional counterpart for a two- and three-dimensional non-linear subspace (blue and green curves). Additionally, the same distance measures are plotted for a two- and three-dimensional linear subspace (red and cyan curves) generated by performing PCA on the autoencoder training data.

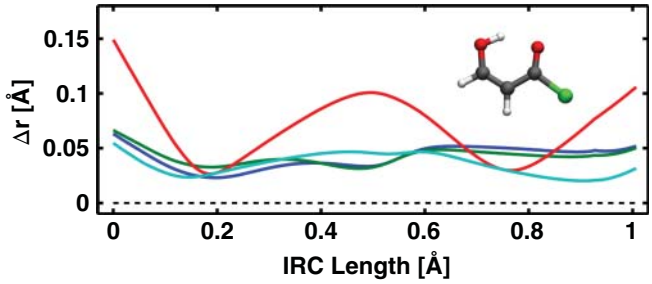


Figure 11.10 Distance between the full-dimensional IRC and an IRC projected onto the low-dimensional subspace. The blue and green curves correspond to a two- and three-dimensional non-linear subspace, respectively, whereas the red and cyan curves belong to a two- and three-dimensional linear subspace, respectively. The linear subspaces were determined by a PCA on the training dataset.

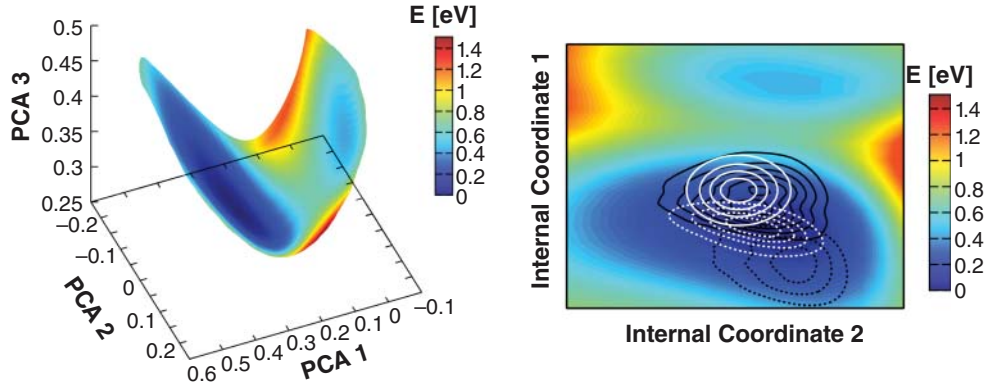


Figure 11.11 Two-dimensional, non-linear PES for the proton transfer in (Z)-hydroxyacryloyl chloride. On the left, the surface is projected onto the first three principal components of the training data. On the right, the same surface is plotted in internal coordinates and overlaid with snapshots of the wave packet propagation at 0 fs (solid white), 17 fs (dashed white), 31 fs (dashed black) and 68 fs (solid black).

It is expected that Δr decreases with the number of additional dimensions as long as these dimensions are beneficial for the description of the reaction. Therefore, Figure 11.10 allows the conclusion that three linear coordinates are needed for an adequate representation of the proton transfer while two non-linear coordinates suffice to achieve the same goal.

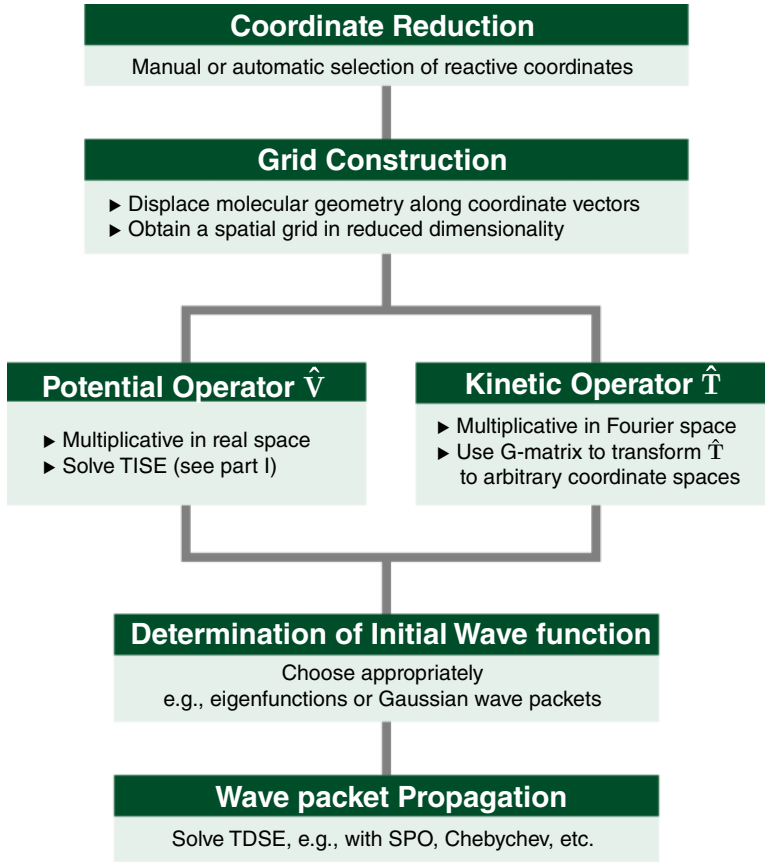
The non-linear PES is shown in Figure 11.11. On the left, the two-dimensional, non-linear surface is projected into a linear coordinate space spanned by the first three principal components of the training data to illustrate its strong curvature. The same surface is shown in internal, non-linear coordinates on the right, where the double minimum structure with the transition state of the reaction on the central barrier is clearly discernible.

A quantum dynamical simulation was performed on this surface for 500 fs with a time-step of 0.24 fs [38]. The Gaussian wave packet was started near the transition state and its temporal evolution is also illustrated in Figure 11.11. During the simulation, the wave packet evolved towards the deeper minimum and was reflected with a cycle period of 70 fs. This is about half of the cycle period of 120 fs reported for the full proton transfer reaction [42]. The timescales of nuclear motion agree with those of the keto-enol switching in malondialdehyde [43] calculated in full dimensionality with semi-classical techniques. The same simulation in linear coordinates results in a too short reaction time. This illustrative example shows that the propagation in non-linear coordinates found

by the autoencoder can adequately describe a chemical reaction and serves as proof-of-concept for this automated procedure. Note, that the interpretation of the wave packet moving in a non-linear coordinate space found by the autoencoder may be challenging.

11.4 Summary and Further Remarks

An overview of the necessary steps to set up a wave packet simulation in reduced coordinates on explicitly evaluated ab intio PESs is given in the following flowchart. All examples discussed in this chapter were performed accordingly.



A detailed description of the individual steps can be found in this chapter (e.g., *G*-matrix, propagators) as well as in chapters 2 to 10 (for the electronic structure methods).

The given examples for arbitrary low-dimensional coordinate spaces are only feasible due to the generality of the *G*-matrix formalism. With this technique the setup of the kinetic operator is straightforward. Its flexibility further allows easy modification of the reactive subspace. For example, structural relaxation can be directly incorporated in the reactive modes [44] as has been demonstrated for the photodisscociation of diphenylmethyl halogenides. Also the reactive subspace can easily be extended with normal modes and their harmonic PESs to test the influence of additional degrees of freedom [27].

The quantum dynamics in reactive subspaces can also be embedded in complex, explicitly treated environments [45–47], allowing to describe the solvent cavity effect during the dissociation or the steric hindrance and electrostatic influence of an RNA strand on the photorelaxation of uracil. In all these examples, subtle changes of the PES were observed to strongly influence the dynamics. In this context, highly local changes in the PES of photoexcited uracil in RNA could severely decrease or prolong the excited state lifetime compared to the isolated base. Alternative methods like MCTDH are able to include more modes at the expense of the topographical information of the PES. The examples mentioned above, however, emphasize the importance of such information for the wave packet behavior and the choice of method depends on the nature of the specific process and the question to be answered.

References

1 Tannor, D.J. (2007). *Introduction to Quantum Mechanics: A Time-Dependent Perspective*. University Science.

2 Marston, C.C. and Balint-Kurti, G.G. (1989). The Fourier grid Hamiltonian method for bound state eigenvalues and eigenfunctions. *J. Chem. Phys.* 91: 3571–3576.

3 Kosloff, R. (1988). Time-dependent quantum-mechanical methods for molecular dynamics. *J. Phys. Chem.* 92: 2087–2100.

4 Feit, M.D. and Fleck, J.A. (1989). Simple spectral method for solving propagation problems in cylindrical geometry with fast Fourier transforms. *Opt. Lett.* 14: 662–664.

5 Kosloff, R. and Kosloff, D. (1983). A Fourier method solution for the time dependent Schrödinger equation: a study of the reaction $H^+ + H_2$, $D^+ + HD$, and $D^+ + H_2$. *J. Chem. Phys.* 79: 1823–1833.

6 Leforestier, C., Bisseling, R.H., Cerjan, C., et al. (1991). A comparison of different propagation schemes for the time dependent Schrödinger equation. *J. Comput. Phys.* 94: 59–80.

7 Tal-Ezer, H. and Kosloff, R. (1984). An accurate and efficient scheme for propagating the time dependent Schrödinger equation. *J. Chem. Phys.* 81: 3967–3971.

8 Kosloff, R. and Tal-Ezer, H. (1986). A direct relaxation method for calculating eigenfunctions and eigenvalues of the Schrödinger equation on a grid. *Chem. Phys. Lett.* 127: 223–230.

9 Franke, R. (1982). Smooth interpolation of scattered data by local thin plate splines. *Comput. Math. Appl.* 8: 273–281.

10 Yarkony, D.R. (2001). Nuclear dynamics near conical intersections in the adiabatic representation: I. The effects of local topography on interstate transitions. *J. Chem. Phys.* 114: 2601–2613.

11 Domcke, W. and Stock, G. (1997). Theory of ultrafast nonadiabatic excited-state processes and their spectroscopic detection in real time. *Adv. Chem. Phys.* 100: 1.

12 Parlant, G. and Yarkony, D.R. (1992). An adiabatic state approach to electronically nonadiabatic wave packet dynamics. *Int. J. Quantum Chem.* 44: 737–749.

13 Hofmann, A. and de Vivie-Riedle, R. (2000). Quantum dynamics of photoexcited cyclohexadiene introducing reactive coordinates. *J. Chem. Phys.* 112: 5054–5059.

14 Hofmann, A. and de Vivie-Riedle, R. (2001). Adiabatic approach for ultrafast quantum dynamics mediated by simultaneously active conical intersections. *Chem. Phys. Lett.* 346: 299–304.

15 de Vivie-Riedle, R. and Hofmann, A. (2004). Nonadiabatic quantum dynamics and control strategies. In: *Conical Intersections: Electronic Structure, Dynamics and Spectroscopy*. World Scientific Publishing.

16 Celani, P., Olivucci, M., Bernard, F. et al. (1994). What happens during the picosecond lifetime of $2A_1$cyclohexa-1, 3-diene? A cas-scf study of the cyclohexadiene/hexatriene photochemical interconversion. *J. Am. Chem. Soc.* 116: 10141–10151.

17 Celani, P., Bernardi, F., Robb, M.A., and Olivucci, M. (1996). Do photochemical ring-openings occur in the spectroscopic state? $^1B_2$ pathways for the cyclohexadiene/hexatriene photochemical interconversion. *J. Phys. Chem.* 100: 19364–19366.

18 Garavelli, M., Bernardi, F., Olivucci, M., et al. (1998). Potential-energy surfaces for ultrafast photochemistry static and dynamic aspects. *Faraday Discuss.* 110: 51–70.

19 Garavelli, M., Bernardi, F., Celani, P. et al. (1998). Minimum energy paths in the excited and ground states of short protonated Schiff bases and of the analogous polyenes. *J. Photochem. Photobiol. A Chem.* 114: 109–116.

20 Berens, P.H. and Wilson, K.R. (1981). Molecular dynamics and spectra. I. Diatomic rotation and vibration. *J. Chem. Phys.* 74: 4872–4882.

21 Stare, J. and Balint-Kurti, G.G. (2003). Fourier grid Hamiltonian method for solving the vibrational Schrödinger equation in internal coordinates: theory and test applications. *J. Phys. Chem. A* 107: 7204–7214.

22 Podolsky, B. (1928). Quantum-mechanically correct form of Hamiltonian function for conservative systems. *Phys. Rev.* 32: 812–816.

23 Zauleck, J.P.P., Thallmair, S. and de Vivie-Riedle, R. (2017). Revisiting an approximation in the Wilson G-matrix formalism and its impact on molecular quantum dynamics. arXiv:1712.10170.

24 Alexandrov, V., Smith, D.M.A., Rostkowska, H., et al. (1998). Theoretical study of the O—H stretching band in 3-hydroxy-2-methyl-4-pyrone. *J. Chem. Phys.* 108: 9685–9693.

25 Eckart, C. (1935). Some studies concerning rotating axes and polyatomic molecules. *Phys. Rev.* 47: 552–558.

26 Dymarsky, A.Y. and Kudin, K.N. (2005). Computation of the pseudorotation matrix to satisfy the Eckart axis conditions. *J. Chem. Phys.* 122: 124103.

27 Keefer, D., Thallmair, S., Matsika, S., and de Vivie-Riedle, R. (2017). Controlling photorelaxation in uracil with shaped laser pulses: a theoretical assessment. *J. Am. Chem. Soc.* 139: 5061–5066.

28 Matsika, S., Spanner, M., Kotur, M., and Weinacht, T.C. (2013). Ultrafast relaxation dynamics of uracil probed via strong field dissociative ionization. *J. Phys. Chem. A* 117: 12796–12801.

29 Zauleck, J.P.P., Thallmair, S., Loipersberger, M., and de Vivie-Riedle, R. (2016). Two new methods to generate internal coordinates for molecular wave packet dynamics in reduced dimensions. *J. Chem. Theory Comput.* 12: 5698–5708.

30 Fukui, K. (1970). Formulation of the reaction coordinate. *J. Phys. Chem.* 74: 4161–4163.

31 García, A.E. (1992). Large-amplitude nonlinear motions in proteins. *Phys. Rev. Lett.* 68: 2696–2699.

32 Abseher, R. and Nilges, M. (2000). Efficient sampling in collective coordinate space. *Proteins Struct. Funct. Genet.* 39: 82–88.

33 Stein, S.A.M., Loccisano, A.E., Firestine, S.M., and Evanseck, J.D. (2006). *Annual Reports in Computational Chemistry. Chapter 13 Principal Components Analysis: A Review of its Application on Molecular Dynamics Data*, vol. 2. Elsevier.

34 David, C.C. and Jacobs, D.J. (2014). Principal component analysis: a method for determining the essential dynamics of proteins. *Methods Mol. Biol.* 1084: 193–226.

35 Isaacson, A.D. and Morokuma, K. (1975). Molecular orbital studies of hydrogen bonds. VIII. Malonaldehyde and symmetric hydrogen bonding in neutral species. *J. Am. Chem. Soc.* 97: 4453–4457.

36 Karlstrom, G., Wennerstrom, H., Jonsson, B., et al. (1975). Intramolecular hydrogen bond. Ab initio MO calculations on the enol tautomer of malondialdehyde. *J. Am. Chem. Soc.* 97: 4188–4192.

37 Ruf, B.A. and Miller, W.H. (1988). A new (cartesian) reaction-path model for dynamics in polyatomic systems, with application to H—atom transfer in malonaldehyde. *J. Chem. Soc. Faraday Trans.* 2, 84: 1523.

38 Zauleck, J.P.P. and de Vivie-Riedle, R. (2018). Constructing grids for molecular quantum dynamics using an autoencoder. *J. Chem. Theory Comput.* 14: 55–62.

39 Rumelhart, D.E., Hinton, G.E., and Williams, R.J. (1986). Learning representations by back-propagating errors. *Nature* 323: 533–536.

40 Hinton, G.E. (2006). Reducing the dimensionality of data with neural networks. *Science* 313: 504–507.

41 Andersen, H.C. (1983). Rattle: a "velocity" version of the shake algorithm for molecular dynamics calculations. *J. Comput. Phys.* 52: 24–34.

42 do N. Varella, M., Arasaki, Y., Ushiyama, H., et al. (2007). Real-time observation of intramolecular proton transfer in the electronic ground state of chloromalonaldehyde: an *ab initio* study of time-resolved photoelectron spectra. *J. Chem. Phys.* 126: 054303.

43 Wolf, K., Mikenda, W., Nusterer, E., and Schwarz, K. (1998). Proton motion in malonaldehyde: an ab initio molecular dynamics study. *J. Mol. Struct.* 448: 201–207.

44 Thallmair, S., Roos, M.K., and de Vivie-Riedle, R. (2016). Design of specially adapted reactive coordinates to economically compute potential and kinetic energy operators including geometry relaxation. *J. Chem. Phys.* 144: 234104.

45 Zauleck, J.P.P., Peschel, M.T., Rott, F. et al. (2018). Ultrafast reactive quantum dynamics coupled to classical solvent dynamics using an Ehrenfest approach. *J. Phys. Chem. A* 122: 2849–2857.

46 Thallmair, S., Zauleck, J.P.P., and de Vivie-Riedle, R. (2015). Quantum dynamics in an explicit solvent environment: a photochemical bond cleavage treated with a combined QD/MD approach. *J. Chem. Theory Comput.* 11: 1987–1995.

47 Reiter, S., Keefer, D., and de Vivie-Riedle, R. (2018). RNA environment is responsible for decreased photostability of uracil. *J. Am. Chem. Soc.* 140: 8714–8720.

12

Multi-Configuration Time-Dependent Hartree Methods: From Quantum to Semiclassical and Quantum-Classical

M. Bonfanti[1], *G. A. Worth*[2], *and I. Burghardt*[1]

[1] *Institute of Physical and Theoretical Chemistry, Goethe University Frankfurt, Max-von-Laue-Str. 7, 60438 Frankfurt, Germany*
[2] *Department of Chemistry, University College London, 20, Gordon St., London, UK*

Abstract

This chapter reviews multi-configurational wave function propagation methods of the Multi-Configuration Time-Dependent Hartree (MCTDH) family, with special emphasis on the connection between MCTDH and its Gaussian wave packet based variants. These include the Gaussian-based G-MCTDH approach and the closely related variational Multi-Configurational Gaussian (vMCG) method, as well as a recently developed two-layered (2L-GMCTDH) version. Besides these fully variational schemes, we also consider a classical-limit formulation leading to a Multi-Configuration Ehrenfest (MCE) type dynamics. This spectrum of methods illustrate the flexibility of variational wave function evolution for high-dimensional quantum dynamics and its semiclassical and quantum-classical variants. A comparative application is presented for ultrafast non-adiabatic dynamics in a charge transfer complex involving up to 100 vibrational modes.

12.1 Introduction

The Multi-Configuration Time-Dependent Hartree (MCTDH) method [1–3] is a powerful approach to the approximate, variational solution of the time-dependent Schrödinger equation (TDSE). In this method, high-dimensional wave functions are represented as sums over configurations – i.e., products of so-called single-particle functions (SPFs) – and the time evolution is broken down into low-dimensional SPF subspaces. Correlations are included by the superposition of configurations, leading to time evolution under multi-configurational mean-field Hamiltonians in the SPF subspaces. This approach allows one to go far beyond the standard method relying on time-independent product basis sets, whose exponential scaling with fN^{f+1} (where f is the number of degrees of freedom and N is the number of basis functions) [3–5] typically restricts propagation to $f \sim 5$. By contrast, MCTDH permits accurate propagation for up to $f \sim 50$ or more degrees of freedom. We refer to Chapter 11 for a detailed discussion of the standard method.

Many applications of the method over the past 25 years have demonstrated the efficiency of MCTDH [3–6]. On the formal side, a number of variants have been introduced, including MCTDH-F for fermionic systems [7–9] and MCTDH-B for bosonic systems [10]. Furthermore, a hierarchical multi-layer version of MCTDH has been developed [11–14], denoted ML-MCTDH,

Quantum Chemistry and Dynamics of Excited States: Methods and Applications,
First Edition. Edited by Leticia González and Roland Lindh.

which is able to take the method towards even higher dimensions, up to hundreds of degrees of freedom. A statistical treatment using density operators was also introduced [15, 16].

In addition, MCTDH has been formulated for moving Gaussian basis sets [17, 18], yielding the G-MCTDH method [19–21] and its variational Multi-Configurational Gaussian (vMCG) variant [22–24], as well as the multi-layer ML-GMCTDH approach [25–27]. This class of methods, even though formulated for Gaussian wave packets (GWPs) that are less flexible than the standard SPFs of the MCTDH method, show good convergence properties and provide a natural connection to more approximate GWP methods and trajectory-based approaches that have a long tradition in semiclassical dynamics [18, 28, 29]. Variational GWP approaches are obviously suitable for system-bath type problems and statistical treatments, as well as for *on-the-fly* dynamics, see the detailed discussion of the direct-dynamics vMCG (DD-vMCG) approach [24] in Chapter 13 as well as the discussion of the Ab Initio Multiple Spawning (AIMS) approach [30] in Chapter 14. These approaches further connect to quantum-classical schemes such as the Multi-Configuration Ehrenfest (MCE) approach [31, 32] which can be understood as a quantum-classical limit of a hybrid SPF/GWP approach within the G-MCTDH framework [33, 34]. See also Chapter 15 for further background on Ehrenfest dynamics.

Complementary to these developments in the chemical physics community, there has been a growing interest in recent years in the connection between MCTDH and low-rank tensor approximation schemes [35, 36] that have been addressed from the mathematical side. In this context, tangent space projection techniques [37] are employed, which introduce a different angle of vision on the variational approach. From this viewpoint, a reformulation of the MCTDH equations of motion (EOM) has recently been proposed [38–40].

Against this background, this chapter spotlights the generality and flexibility of variational wave packet approaches and their applications, with special emphasis on the connection between MCTDH and its GWP based variants. To summarize our perspective, Figure 12.1 illustrates the relations between the approaches that are discussed in this chapter.

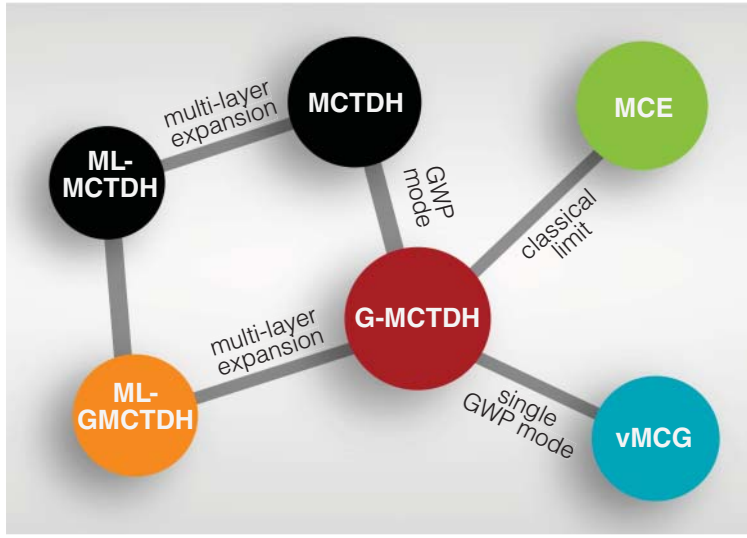


Figure 12.1 Schematic illustration of the MCTDH related methods that are discussed in this chapter. See text for the acronyms used in the figure.

As an explanatory note regarding GWP based methods, we use the term *semiclassical* to refer to the fact that GWP bases preserve the classical phase space structures in a quantum context. In this sense, the hybrid G-MCTDH method has been termed a *quantum-semiclassical* method [19]. In contrast, the MCE approach is referred to as a *mixed quantum-classical* approach – i.e., the quantum-classical limit of the G-MCTDH formulation in the strict sense of Ref. [33]. However, MCE also exemplifies that these boundaries are loosely defined, since several alternative versions of the MCE approach [31, 32, 34], as also discussed below, rather fall into a quantum-semiclassical regime.

The remainder of the chapter is organized as follows. In the second section, we introduce the MCTDH equations of motion and connect to the above mentioned tangent space projection techniques and low-rank tensor approximation schemes. We further introduce the hierarchically structured ML-MCTDH variant. The third section is specifically devoted to GWP based methods, notably the G-MCTDH, vMCG and ML-GMCTDH approaches. The fourth section then addresses the quantum-classical limit and the resulting MCE dynamics. In the fifth section, a brief How-To summarizes some practical aspects. Finally, the sixth section illustrates the various levels of treatment for a case study of ultrafast non-adiabatic dynamics in a donor–acceptor complex [41, 42], and the last section is the conclusion.

12.2 Time-Dependent Variational Principle and MCTDH

In this section, we briefly review the MCTDH method and its derivation from the time-dependent Dirac–Frenkel Variational Principle (DFVP). We also address the basic concept of the hierarchical ML-MCTDH approach.

12.2.1 Variational Principle and Tangent Space Projections

The DFVP is the overarching principle of MCTDH and its variants, by defining the optimal time evolution of the approximate wave functions that characterize the different approaches shown in Figure 12.1. Specifically, the MCTDH equations of motion are derived from a multi-configurational *ansatz* for the wave function – i.e., a sum over Hartree products as detailed in Section 12.2.2 – in conjunction with the DFVP [43–48],

$$\langle \delta\Psi | i\partial_t - \hat{H} | \Psi \rangle = 0 \tag{12.1}$$

where we use $\partial_t = \partial/\partial t$ and units such that $\hbar = m = 1$. This convention will be employed throughout this chapter. In Eq. (12.1), $\delta\Psi$ is an allowed variation of the wave function which keeps the form of the chosen *ansatz*. The linear variation $\delta\Psi$ is an element of the tangent space $T_\Psi\mathcal{M}$ (see Figure 12.2) which is defined with respect to the smooth submanifold $\mathcal{M}$ of the Hilbert space where we are seeking an approximation to the time evolving quantum state Ψ. The time evolution resulting from the DFVP according to Eq. (12.1) guarantees that Ψ is the best possible approximation to the exact solution of the time-dependent Schrödinger equation, given the restricted form of the wave function *ansatz* [37].

The DFVP has been shown to be equivalent to the McLachlan Variational Principle (MLVP) [47], provided that the tangent space $T_\Psi\mathcal{M}$ is a complex linear space [46, 94]. The MLVP formulates a geometric condition in terms of which the approximate time derivative of the wave function features a minimal distance from the exact solution,

$$\delta \| i\dot{\Psi} - \hat{H}\Psi \|^2 = 0 \tag{12.2}$$

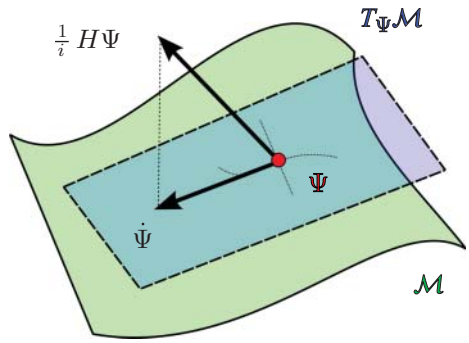


Figure 12.2 Pictorial representation of the tangent-space projection. At a given point Ψ of the variational manifold $\mathcal{M}$ of the full Hilbert space, the tangent-space $T_\Psi\mathcal{M}$ is constructed as the vector space spanned by the first-order variation of the wave function. According to the DFVP, the best local approximation to the time-derivative $(1/i)\,\hat{H}\,\Psi$ is given by $\dot{\Psi}$, its orthogonal projection onto the tangent space. Reprinted from Ref. [40] with permission from Elsevier.

This condition can be naturally fulfilled by introducing tangent space projectors [37] such that the best approximate wave function derivative, at any point in time, is constructed as the orthogonal projection of $\hat{H}\Psi$ onto the tangent space $T_\Psi\mathcal{M}$,

$$i\dot{\Psi} = \hat{\mathcal{P}}(\Psi)\hat{H}\Psi \tag{12.3}$$

where $\hat{\mathcal{P}}(\Psi)$ is the orthogonal projector onto the tangent space. Given that $\dot{\Psi}$ is a vector of $T_\Psi\mathcal{M}$ by construction, it belongs to the range of the projector $\hat{\mathcal{P}}(\Psi)$. Thus Eq. (12.3) can be rearranged as a projected TDSE [37],

$$\hat{\mathcal{P}}(\Psi)[i\dot{\Psi} - \hat{H}\Psi] = 0 \tag{12.4}$$

A schematic representation of the tangent space projection is shown in Figure 12.2. Details of this formulation of the DFVP, which is most common in the mathematical literature, can be found in Ref. [37].

In the standard derivation of the MCTDH method, Eq. (12.1) is employed to obtain the EOMs. However, we will also discuss Eq. (12.4), which opens an alternative perspective on the MCTDH equations. Notably, this route has recently been employed in deriving a novel projector splitting formulation of the MCTDH equations [38–40], i.e., relying on a suitable splitting of $\hat{\mathcal{P}}(\Psi)$.

12.2.2 MCTDH: Variational Multi-Configurational Wave Functions

Since the MCTDH approach, in its original form, defines the roadmap to be followed in the context of the various related schemes to be discussed in this chapter, we start by reviewing the MCTDH wave function form and the derivation of the variational EOMs.

12.2.2.1 MCTDH Wave Function *Ansatz*

The MCTDH *ansatz* [1–3] corresponds to a multi-configurational form of the wave function [3, 37],

$$\begin{aligned}\Psi(\mathbf{x},t) &= \sum_J A_J(t)\Phi_J(\mathbf{x},t)\\ &= \sum_{j_1=1}^{n_1}\sum_{j_2=1}^{n_2}\cdots\sum_{j_f=1}^{n_f} A_{j_1 j_2,\ldots j_f}(t)\prod_{\kappa=1}^{f}\varphi_{j_\kappa}^{(\kappa)}(\mathbf{x}_\kappa,t)\,.\end{aligned} \tag{12.5}$$

with complex-valued coefficients $A_J \equiv A_{j_1 j_2,\ldots j_f}$ and configurations $\Phi_J \equiv \prod_{\kappa=1}^{f}\varphi_{j_\kappa}^{(\kappa)}$ which are Hartree products of time-dependent single-particle functions (SPFs) $\varphi_j^{(\kappa)}$. Here, the multi-index

$J = (j_1, j_2, \dots, j_f)$ collects all SPF indices for a given configuration Φ_J. The index $\kappa = 1, 2, \dots, f$ enumerates a collection of modes $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_f)$, where each mode $\mathbf{x}_\kappa$ corresponds either to a single degree of freedom or else a combination of degrees of freedom $\mathbf{x}_\kappa = (x_{\kappa,1}, \dots, x_{\kappa,d_\kappa})$. In the latter case, $\mathbf{x}_\kappa$ represents a so-called combined mode ($d_\kappa > 1$) [3, 49]. For each κ, n_κ SPFs $\varphi_{j_\kappa}^{(\kappa)}$, with $j_\kappa = 1, 2, \dots, n_\kappa$, are specified; i.e., n_κ defines the dimension of the SPF basis for a given mode κ.

Throughout this chapter, small letters x_i are used to label nuclear coordinates, mostly corresponding to internal vibrational modes; also, we occasionally employ small letters r_i or capital letters R_i to label specific internal coordinates. (By contrast, electronic coordinates do not appear explicitly, i.e., Dirac notation is used to represent electronic states, generally referring to a diabatic picture.)

When expressing the time-evolving SPFs in a time-independent product basis $\{\chi_{i_1}^{(1)} \dots \chi_{i_f}^{(f)}\}$ (sometimes denoted primitive basis), Eq. (12.5) can be re-expressed as

$$\Psi(\mathbf{x}, t) = \sum_{i_1=1}^{N_1} \sum_{i_2=1}^{N_2} \dots \sum_{i_f=1}^{N_f} Y_{i_1,i_2,\dots i_f}(t) \prod_{\kappa=1}^{f} \chi_{i_\kappa}^{(\kappa)}(\mathbf{x}_\kappa) \tag{12.6}$$

with

$$Y_{i_1,i_2,\dots i_f}(t) = \sum_{j_1=1}^{n_1} \dots \sum_{j_f=1}^{n_f} A_{j_1 j_2 \dots j_f}(t) \prod_{\kappa}^{f} U_{i_\kappa j_\kappa}^{(\kappa)}(t) \tag{12.7}$$

where $U_{i_\kappa j_\kappa}^{(\kappa)} = \langle \chi_{i_\kappa}^{(\kappa)} | \varphi_{j_\kappa}^{(\kappa)} \rangle$ is the representation matrix of the κ-mode SPFs on the primitive grid. From a tensor algebra perspective [50], Eq. (12.7) is known as Tucker decomposition of the tensor $Y_{i_1,i_2,\dots i_f}$ into the core tensor $A_{j_1 j_2 \dots j_f}$ and the set of matrices $U_{i_\kappa j_\kappa}^{(\kappa)}$. As the number of SPFs is obviously smaller than the size of the primitive basis, the Tucker decomposition entails a reduction in dimensionality of the original tensor, taking advantage of its possible sparsity.

The representation of Eq. (12.5) is not unique, since an alternative SPF basis can be found that spans the κth subspaces such that

$$\Psi = \sum_J A_J \prod_{\kappa=1}^{f} \varphi_{j_\kappa}^{(\kappa)} = \sum_{\bar{J}} \tilde{A}_{\bar{J}} \prod_{\kappa=1}^{f} \tilde{\varphi}_{\bar{j}_\kappa} \ ,$$

This gauge freedom in defining the MCTDH *ansatz* is exploited by specifying initial SPFs $\varphi_j^{(\kappa)}(0)$ that are orthonormal and unambiguously fixing their time evolution such that they stay orthonormal at all times [3],

$$\langle \varphi_j^{(\kappa)}(t) | \varphi_{j'}^{(\kappa)}(t) \rangle = \delta_{jj'} .$$

The remaining freedom in the definition of the SPFs is captured by an explicit, unitary time evolution inside the κth subspaces,

$$i \langle \varphi_l^{(\kappa)} | \dot{\varphi}_j^{(\kappa)} \rangle - \langle \varphi_l^{(\kappa)} | \hat{h}_c^{(\kappa)} | \varphi_j^{(\kappa)} \rangle = 0 \tag{12.8}$$

or equivalently,

$$\hat{P}^{(\kappa)}(t)(i\partial_t - \hat{h}_c^{(\kappa)}) | \varphi_j^{(\kappa)}(t) \rangle = |0\rangle \tag{12.9}$$

where

$$\hat{P}^{(\kappa)}(t) = \sum_j |\varphi_j^{(\kappa)}(t)\rangle \langle \varphi_j^{(\kappa)}(t)| \tag{12.10}$$

is the orthogonal projector onto the κth subspace. In Eqs. (12.8)–(12.9), the so-called constraint operator $\hat{h}_c^{(\kappa)}$ is an arbitrary Hermitian operator; this operator is most often set to zero [3], as is also the case in this chapter. Note that the projector $\hat{P}^{(\kappa)}$ is different from the tangent space projector

$\widehat{\mathcal{P}}(\Psi)$ of Eq. (12.3), which is represented in calligraphic style for clarity. With the gauge Eq. (12.8), one obtains unambiguous EOMs for the coefficients A_J and SPFs $\varphi_j^{(\kappa)}$.

As a notational issue, the wave function can be expressed as a product of SPFs and single-hole functions (SHFs) denoted $\psi_{j_\kappa}^{(\kappa)}$ [3],

$$\Psi(\mathbf{x},t) = \sum_{j_\kappa=1}^{n_\kappa} \varphi_{j_\kappa}^{(\kappa)}(\mathbf{x}_\kappa,t)\psi_{j_\kappa}^{(\kappa)}(\mathbf{x}_1,\dots,\mathbf{x}_{\kappa-1},\mathbf{x}_{\kappa+1},\dots,\mathbf{x}_N,t) \tag{12.11}$$

which is convenient when defining the evolution equations for the κth subspace. The SHFs are non-orthogonal – despite the orthogonality of the SPFs – and their overlap appears as the reduced density matrix in the κth subspace,

$$\rho_{l_\kappa l'_\kappa}^{(\kappa)} = \langle \psi_{l_\kappa}^{(\kappa)} | \psi_{l'_\kappa}^{(\kappa)} \rangle \tag{12.12}$$

noting that $\rho_{l_\kappa l'_\kappa}^{(\kappa)} = \langle \varphi_{l'_\kappa}^{(\kappa)} | \widehat{\rho}^{(\kappa)} | \varphi_{l_\kappa}^{(\kappa)} \rangle$ where $\widehat{\rho}^{(\kappa)} = \mathrm{Tr}_{\kappa'\neq\kappa}\{\Psi\Psi^*\}$ is the reduced density operator in the κth subspace.

12.2.2.2 MCTDH Equations of Motion

The variational EOMs for the wave function *ansatz* Eq. (12.5) can be derived from the DFVP in the form of Eq. (12.1), or alternatively, from Eq. (12.4). The former route corresponds to the standard procedure, where the wave function variation is specified as

$$\delta\Psi = \sum_J \delta A_J \, \Phi_J + \sum_\kappa \left(\sum_{l_\kappa} \delta\varphi_{l_\kappa}^{(\kappa)} \psi_{l_\kappa}^{(\kappa)} \right) \tag{12.13}$$

and independent variations δA_J and $\delta\varphi_{l_\kappa}^{(\kappa)}$ are inserted into Eq. (12.1).

Alternatively, the tangent space projector of Eq. (12.3) can be obtained explicitly [39, 40],

$$\widehat{\mathcal{P}}(\Psi) = \widehat{\mathcal{P}}_0(\Psi) + \sum_\kappa \widehat{\mathcal{P}}_\kappa(\Psi) \tag{12.14}$$

and the component projectors pertaining to the $\boldsymbol{A}$ coefficients and SPFs are inserted into the projected TDSE of Eq. (12.4).

In Eq. (12.14), the projectors $\widehat{\mathcal{P}}_0(\Psi)$ and $\widehat{\mathcal{P}}_\kappa(\Psi)$ correspond to a projector on the configurations $|\Phi_J\rangle$ in the full space, and a combination of SPF subspace projectors $\widehat{P}^{(\kappa)}$ (see Eq. (12.10)) and SHF projectors $\widehat{Q}^{(\kappa)}$,

$$\widehat{\mathcal{P}}_0(\Psi) = \sum_J |\Phi_J\rangle\langle\Phi_J| \qquad \widehat{\mathcal{P}}_\kappa(\Psi) = (\widehat{1} - \widehat{P}^{(\kappa)}) \otimes \widehat{Q}^{(\kappa)} \tag{12.15}$$

where the SHF projector is given as

$$\widehat{Q}^{(\kappa)} = \sum_{l,l'} |\psi_l^{(\kappa)}\rangle \, (\rho^{(\kappa)})^{-1}_{ll'} \, \langle \psi_{l'}^{(\kappa)}| \tag{12.16}$$

noting that the form of $\widehat{Q}^{(\kappa)}$ follows from the non-orthogonality of the SHFs. Reference [40] provides further details of this procedure.

The EOMs resulting from either procedure – i.e., Eq. (12.13) in conjunction with Eq. (12.1), or else Eq. (12.14) in conjunction with Eq. (12.4) – are given as follows:

$$i\dot{A}_I = \sum_J \langle \Phi_I | \widehat{H} | \Phi_J \rangle \, A_J \tag{12.17}$$

where $H_{IJ} = \langle \Phi_I | \widehat{H} | \Phi_J \rangle$ corresponds to elements of the Hamiltonian matrix in the basis of configurations, and

$$i\dot{\boldsymbol{\varphi}}^{(\kappa)} = (\widehat{1} - \widehat{P}^{(\kappa)})(\boldsymbol{\rho}^{(\kappa)})^{-1} \, \widehat{\boldsymbol{H}}^{(\kappa)} \, \boldsymbol{\varphi}^{(\kappa)} \tag{12.18}$$

where $\boldsymbol{\varphi}^{(\kappa)}$ is the vector composed of the SPFs for mode κ, $\hat{P}^{(\kappa)}$ is the subspace projector of Eq. (12.10), $\boldsymbol{\rho}^{(\kappa)}$ is the reduced density matrix of Eq. (12.12), and $\hat{\boldsymbol{H}}^{(\kappa)}$ is a matrix of mean-field Hamiltonian operators,

$$\hat{H}^{(\kappa)}_{jk} = \langle \psi^{(\kappa)}_j | \hat{H} | \psi^{(\kappa)}_k \rangle \tag{12.19}$$

where $|\psi^{(\kappa)}_k\rangle$ are the SHFs of Eq. (12.11). Hence, besides the standard equations for the wave function coefficients, the SPFs are found to be coupled by mean fields: i.e., every SPF $\varphi^{(\kappa)}_j$ evolves under an effective single-particle Hamiltonian obtained by averaging the full many-particle Hamiltonian $\hat{H}$ over the remaining degrees of freedom $\kappa' \neq \kappa$. In Eq. (12.18), the $(\hat{1} - \hat{P}^{(\kappa)})$ projection results from the tangent space projection $\hat{\mathcal{P}}_\kappa$ of Eq. (12.15) and implies that the SPF evolution exclusively takes into account the part of the time evolution that is not absorbed into Eq. (12.17) for the $\boldsymbol{A}$ coefficients.

In the simplest, single-configurational Time-Dependent Hartree (TDH) case, a single effective mean-field Hamiltonian is given per mode κ. In the multi-configurational (MCTDH) case, the SPFs of each mode κ evolve under a matrix of mean-field Hamiltonians.

Besides the numerical effort for the propagation of the coefficient vector $\boldsymbol{A}$, the integrations involved in the calculation of the mean field Hamiltonian are the numerically most expensive parts of MCTDH. To carry out the relevant integrals, the SPFs $\varphi^{(\kappa)}_j$, which are of arbitrary form, are expanded in a primitive time-independent basis set $\{\chi^{(1)}_{i_1} \dots \chi^{(f)}_{i_f}\}$ (see Eq. (12.6)), usually using a Discrete Variable Representation (DVR) [3]. Furthermore, a sum-over-products (SOP) form of the Hamiltonian is required for an efficient integral evaluation, matching the product form of the wave function.

In Eq. (12.18), the presence of the inverse of the subspace density matrix, $(\boldsymbol{\rho}^{(\kappa)})^{-1}$, renders the EOMs highly non-linear. At the start of the propagation, singularities typically occur since most SPFs are initially unoccupied, necessitating a numerical regularization procedure [3]. An improved approach has been developed in Ref. [51] where optimal initial SPFs are constructed. Also, Ref. [52] proposes a novel regularization scheme for the $\boldsymbol{A}$ coefficient tensor. Finally, the recently developed projector splitting algorithm [38–40] circumvents the direct inversion of the density matrix, by recasting the EOMs in a linear form, using an auxiliary set of orthogonal SHFs.

12.2.3 ML-MCTDH: Hierarchical Representations

While MCTDH was originally employed for one-dimensional SPFs ($d_\kappa = 1$, see Section 12.2.2.1), the use of multi-dimensional SPFs ($d_\kappa > 1$) – as foreseen in Eq. (12.5) – can be advantageous to accommodate correlations within the combined-mode subspaces [3, 49]. This strategy allows one to substantially reduce the number of configurations, i.e., the length of the $\boldsymbol{A}$ vector. Mode combination was therefore a key step towards treating high-dimensional systems [49, 53]. However, a balance needs to be found between the gain in the $\boldsymbol{A}$ vector propagation and the cost of propagating combined-mode SPFs, which scales exponentially with d_κ [3].

In view of this, an improved and systematic strategy is to represent each combined mode by an MCTDH type *ansatz* for a given subspace. This leads to the following multi-layer MCTDH (ML-MCTDH) *ansatz* [11–14], where the hierarchy is started by expanding the overall wave function Ψ in first-layer SPFs:

$$\begin{aligned}\Psi(t) &= \sum_J A^{[1]}_J(t)\Phi^{[1]}_J(t) \\ &= \sum_J A^{[1]}_J(t) \prod_{\kappa_1=1}^{f^{[1]}} \varphi^{[1](\kappa_1)}_{j_{\kappa_1}}(t)\end{aligned} \tag{12.20}$$

While this expansion is formally identical to Eq. (12.5), the *same* type of multi-configurational representation is now also used for the first-layer SPFs $\varphi_{j_{\kappa_1}}^{[1](\kappa_1)}$, which are in turn expanded in second-layer SPFs $\varphi_{j_{\kappa_2}}^{[2](\kappa_1,\kappa_2)}$:

$$\begin{aligned}\varphi_j^{[1](\kappa_1)}(t) &= \sum_J A_{j,J}^{[2](\kappa_1)} \Phi_J^{[2](\kappa_1)}(t) \\ &= \sum_J A_{j,J}^{[2](\kappa_1)} \prod_{\kappa_2=1}^{f_{\kappa_1}^{[2]}} \varphi_{j_{\kappa_2}}^{[2](\kappa_1,\kappa_2)}(t)\end{aligned} \tag{12.21}$$

Moreover, the procedure can be re-iterated for an M-layer representation, such that the SPFs of the first $M-1$ layers are given as,

$$\begin{aligned}\varphi_j^{[m-1](\kappa_1\dots\kappa_{m-1})}(t) &= \sum_J A_{j,J}^{[m](\kappa_1\dots\kappa_{m-1})}(t) \Phi_J^{[m](\kappa_1\dots\kappa_{m-1})}(t) \\ &= \sum_J A_{j,J}^{[m](\kappa_1\dots\kappa_{m-1})}(t) \prod_{\kappa_m=1}^{f_{\kappa_1\dots\kappa_{m-1}}^{[m]}} \varphi_{j_{\kappa_m}}^{[m](\kappa_1\dots\kappa_m)}(t)\end{aligned} \tag{12.22}$$

where $m = \{2, 3, \dots, M\}$ runs over the number of layers and the SPF index $\kappa_m = 1, 2, \dots, f_{\kappa_1\dots\kappa_m}^{[m]}$ runs over the mth-layer modes. Finally, the SPFs of the final (Mth) layer are represented in a primitive DVR basis.

From a mathematical tensor algebra perspective, the representation of Eqs. (12.20)–(12.22) is also denoted hierarchical Tucker decomposition [50], in line with the discussion in Section 12.2.2.1.

The EOMs for the multi-layer approach involve a hierarchy of mth-layer coefficients evolving under the corresponding mth-layer mean fields. For details, we refer to Refs. [11, 12] and Refs. [13, 14] which describe a recursive algorithm to represent these equations. Below, we will focus on a two-layer and multi-layer variant of Gaussian-based MCTDH which shares the basic features of the above representation.

12.3 Gaussian-Based MCTDH

We now turn to a variant of the MCTDH scheme which employs parametrized SPFs, instead of the conventional SPFs which rely on the representation in a primitive basis. Specifically, GWPs are highly suitable for this approach, given that Gaussian matrix elements can be straightforwardly calculated analytically. Hence, we will be concerned with the Gaussian-based G-MCTDH method [19–21], as well as the related vMCG [22–24] and ML-GMCTDH [25–27] approaches.

12.3.1 G-MCTDH and vMCG

In the G-MCTDH approach [19–21], some or all SPFs are represented by GWPs, whose time evolution is defined in terms of the time-evolving GWP center in phase-space $(\boldsymbol{q}, \boldsymbol{p})$, along with the time-evolving width and phase. While some flexibility in the wave function representation is lost, the potential gain is an efficient treatment of a large number of modes, benefiting from the analytical calculation of high-dimensional integrals. Applications of the G-MCTDH method range from model Hamiltonians for dissipative dynamics [20] to high-dimensional vibronic coupling dynamics [21], and the realistic description of photoexcited chromophores in rare-gas environments [55, 56]. Also, G-MCTDH wave functions were recently combined with a mixed-state

density matrix representation [57]. In the context of *on-the-fly* applications, the vMCG variant of the G-MCTDH scheme has mainly been employed [23, 24], where all SPFs are replaced by GWPs in the full dimensionality, providing a semiclassical, localized, variational basis that is superior to classically evolving GWPs.

12.3.1.1 G-MCTDH Wave Function Ansatz

In Ref. [19], hybrid configurations were introduced which restrict a subset of modes in the κth subspace to Gaussian form. This ansatz was specifically tailored to system-bath type problems [20, 21] where a subset of "primary" degrees of freedom are represented by fully flexible SPFs $\varphi_j^{(\kappa)}$ while the remaining – potentially many – "secondary" degrees of freedom are combined into less flexible GWPs $g_j^{(\kappa)}$,

$$\Psi = \sum_J A_J \left(\prod_{\kappa=1}^{p} \varphi_{j_\kappa}^{(\kappa)} \right) \left(\prod_{\kappa=p+1}^{f} g_{j_\kappa}^{(\kappa)} \right) \tag{12.23}$$

with $\varphi_j^{(\kappa)}$ as in standard MCTDH and

$$g_j^{(\kappa)}(\mathbf{x}_\kappa, t) = \exp[\mathbf{x}_\kappa^T \boldsymbol{a}_j^{(\kappa)}(t)\, \mathbf{x}_\kappa + (\boldsymbol{\xi}_j^{(\kappa)}(t))^T \mathbf{x}_\kappa + \eta_j^{(\kappa)}(t)] \tag{12.24}$$

multi-dimensional GWPs whose time evolution is determined by the complex time-dependent parameters $\boldsymbol{\Lambda}_j^{(\kappa)}(t) = (\boldsymbol{a}_j^{(\kappa)}(t), \boldsymbol{\xi}_j^{(\kappa)}(t), \eta_j^{(\kappa)}(t))$. Here, the complex symmetric, negative definite, matrix $\boldsymbol{a}_j^{(\kappa)}$ controls the width of the Gaussian $g_j^{(\kappa)}$ and its off-diagonal elements describe correlations between the degrees of freedom within a GWP combined mode. The complex vector $\boldsymbol{\xi}_j^{(\kappa)}$ describes the center of the GWP in phase space. (More precisely, writing $\boldsymbol{\xi}_j^{(\kappa)} = -2\boldsymbol{a}_j^{(\kappa)}\boldsymbol{q}_j^{(\kappa)} + i\boldsymbol{p}_j^{(\kappa)}$ with $\boldsymbol{q}_j^{(\kappa)}$ and $\boldsymbol{p}_j^{(\kappa)}$ real, $\boldsymbol{q}_j^{(\kappa)}$ is the GWP center in position space and $\boldsymbol{p}_j^{(\kappa)}$ is its center in momentum space.) Finally, the real part of $\eta_j^{(\kappa)}$ fixes the GWP norm $\|g_j^{(\kappa)}\|$ and the imaginary part of $\eta_j^{(\kappa)}$ is a phase factor.

Starting from Eq. (12.24), two types of GWPs are used in practice: thawed Gaussians (TGs) [58–60] whose width matrix $\boldsymbol{a}_j^{(\kappa)}$ is a time-dependent variational parameter and frozen Gaussians (FGs) [29] whose width matrix $\boldsymbol{a}_j^{(\kappa)}$ is kept fixed, such that the parameter vector is reduced to $\boldsymbol{\Lambda}_j^{(\kappa)}(t) = (\boldsymbol{\xi}_j^{(\kappa)}(t), \eta_j^{(\kappa)}(t))$. Like the general multi-configurational form Eq. (12.5), the representation Eq. (12.23) of the wave function Ψ is not unique. At the level of the Gaussians $g_j^{(\kappa)}$ this results in a free choice of $\eta_j^{(\kappa)}(t)$. (In fact, the variational equations for this parameter are ill determined [19].) Conventionally, the real part $\mathrm{Re}(\eta_j^{(\kappa)}(t))$ is fixed such that $g_j^{(\kappa)}(t)$ is always normalized, and the imaginary part $\mathrm{Im}(\eta_j^{(\kappa)}(t))$ is either set to zero [20, 21] or else is taken to evolve with the classical action [61, 62]. Thus the parameters whose time evolution is determined by the DFVP Eq. (12.1) reduce to $\boldsymbol{\Lambda}_j^{(\kappa)}(t) = (\boldsymbol{a}_j^{(\kappa)}(t), \boldsymbol{\xi}_j^{(\kappa)}(t))$ for TGs and $\boldsymbol{\Lambda}_j^{(\kappa)}(t) = \boldsymbol{\xi}_j^{(\kappa)}(t)$ for FGs. Note that the free choice of $\eta(t)$ does not suffice to make the Gaussians $g_j^{(\kappa)}$ mutually orthogonal, i.e., they always form a non-orthogonal basis set.

In the majority of applications, multi-dimensional FGs are used, in view of the numerical instabilities incurred by TG propagation. (See, however, Ref. [20] for an example of TG-based G-MCTDH propagation.) With a constant diagonal width matrix, $(\boldsymbol{a}_j^{(\kappa)})_{kl} = \delta_{kl}(\boldsymbol{a}_j^{(\kappa)})_{kk}$, and the phase convention $\mathrm{Im}(\eta_j^{(\kappa)}(t)) = 0$, Eq. (12.24) reduces to

$$g_j^{(\kappa)}(\mathbf{x}_\kappa, t) = \prod_{l=1}^{d_\kappa} N((\boldsymbol{a}_j^{(\kappa)})_{ll}, (\boldsymbol{\xi}_j^{(\kappa)})_l(t)) \exp[(\boldsymbol{a}_j^{(\kappa)})_{ll} x_{\kappa,l}^2 + (\boldsymbol{\xi}_j^{(\kappa)})_l(t) x_{\kappa,l}] \tag{12.25}$$

with $N(a,\xi) = (-2\mathrm{Re}(a)/\pi)^{1/4}\exp((\mathrm{Re}(\xi))^2/4\mathrm{Re}(a))$ and the shorthand notation $x_{\kappa,l} = (\mathbf{x}_\kappa)_l$. That is, each multi-dimensional FG corresponds to a product of one-dimensional FGs, or coherent states [61–64].

Finally, the ansatz Eq. (12.23) is often restricted to a "GWP-only" wave function where no primary modes are present ($p = 0$),

$$\Psi = \sum_J A_J \prod_{\kappa=1}^{f} g_{j_\kappa}^{(\kappa)}(\mathbf{x}_\kappa, t) \tag{12.26}$$

A special case of this restricted form is employed in the variational Multi-Configurational Gaussian (vMCG) scheme [22, 23, 65, 66] where all degrees of freedom are combined into a single high-dimensional GWP ($p = 0, f = 1$),

$$\Psi_{\mathrm{vMCG}} = \sum_j A_j(t) g_j(\mathbf{x}, t) \tag{12.27}$$

Assuming that Eq. (12.26) and Eq. (12.27) employ the FG type GWPs of Eq. (12.25), both wave functions can be broken down into products of one-dimensional FGs. The difference lies in the number of coefficients, noting that in Eq. (12.26), $A_J \equiv A_{j_1 j_2,\ldots j_f}$ is a multi-index as in the original MCTDH wave function form of Eq. (12.5). In contrast, vMCG does not employ an MCTDH type structure of the wave function – i.e., superpositions of Hartree products with coefficients labeled by multi-indices – such that the coefficients A_j of Eq. (12.27) are single-index quantities. Hence, vMCG has the appeal of using the simplest GWP based wave function form, but does not benefit from the reduction of the dynamical problem to low-dimensional subspaces, as further discussed in Section 12.3.1.3.

12.3.1.2 G-MCTDH Equations of Motion

The derivation of the G-MCTDH equations of motion proceeds analogously to the MCTDH case, but the wave function variation now includes terms that relate to the GWP parameters. A detailed description of the derivation is given in Ref. [19]. In the general hybrid case of Eq. (12.23), the G-MCTDH equations of motion combine a MCTDH-like evolution in the primary subspace with the GWP evolution in the Gaussian subspaces. Differences from the MCTDH equations will appear due to the non-orthogonality of the GWP basis function, and hence, the configurations of Eq. (12.23).

In detail, the equations read as follows [19] for the time-dependent $\boldsymbol{A}$ coefficients,

$$i\boldsymbol{S}\dot{\boldsymbol{A}} = (\boldsymbol{H} - i\boldsymbol{\tau})\boldsymbol{A} \tag{12.28}$$

time-dependent primary-mode SPFs,

$$i\dot{\boldsymbol{\varphi}}^{(\kappa)} = (\hat{1} - \hat{P}^{(\kappa)})(\boldsymbol{\rho}^{(\kappa)})^{-1}\ \langle \hat{H}\rangle^{(\kappa)}\ \boldsymbol{\varphi}^{(\kappa)} \tag{12.29}$$

and time-dependent GWP parameters $\boldsymbol{\Lambda}$,

$$i\boldsymbol{C}^{(\kappa)}\dot{\boldsymbol{\Lambda}}^{(\kappa)} = \boldsymbol{Y}^{(\kappa)}\ . \tag{12.30}$$

where the matrix $\boldsymbol{C}^{(\kappa)}$ and vector $\boldsymbol{Y}^{(\kappa)}$ are given as follows,

$$\begin{aligned} C^{(\kappa)}_{j\alpha j'\beta} &= \rho^{(\kappa)}_{jj'}\langle \partial_\alpha g^{(\kappa)}_j | (1 - \hat{P}^{(\kappa)}) | \partial_\beta g^{(\kappa)}_{j'}\rangle \\ Y^{(\kappa)}_{j\alpha} &= \sum_{j'} \langle \partial_\alpha g^{(\kappa)}_j | (1 - \hat{P}^{(\kappa)}) \hat{H}^{(\kappa)}_{jj'} | g^{(\kappa)}_{j'}\rangle \end{aligned} \tag{12.31}$$

In Eq. (12.28), the Hamiltonian matrix elements $H_{JJ'} = \langle \Phi_J | \widehat{H} | \Phi_{J'} \rangle$ refer to the basis of configurations, as in Eq. (12.17). However, the overlap matrix,

$$S_{JJ'} = \langle \Phi_J | \Phi_{J'} \rangle = \prod_{\kappa=1}^{p} \delta_{j_\kappa l_\kappa} \prod_{\kappa=p+1}^{f} S^{(\kappa)}_{j_\kappa l_\kappa} \tag{12.32}$$

now appears in the coefficient equations Eq. (12.28), along with the differential overlap $\tau_{JJ'} = \langle \Phi_J | \partial_t \Phi_{J'} \rangle$, due to the non-orthogonality of the GWP basis. Similar EOMs for the time-evolving coefficients appear in various GWP-based methods, including the AIMS approach [30] and the MCE approach [31, 34]. These approaches are further discussed in Chapters 14 and 15, respectively.

In Eq. (12.31), $\partial_\alpha g^{(\kappa)}_j = (\partial g^{(\kappa)}_j / \partial \Lambda^{(\kappa)}_{j\alpha})$ denotes the partial derivatives with respect to the GWP parameters, and $\widehat{H}^{(\kappa)}_{jj'}$ represent the mean-field Hamiltonian operator as in Eq. (12.19). Further, $\widehat{P}^{(\kappa)}$ is the orthogonal projector on the κth subspace, which in the case of a non-orthogonal basis set is given by

$$\widehat{P}^{(\kappa)} = \sum_{j,j'} |g^{(\kappa)}_j\rangle ((\boldsymbol{S}^{(\kappa)})^{-1})_{jj'} \langle g^{(\kappa)}_{j'}| \tag{12.33}$$

with $S^{(\kappa)}_{jj'} = \langle g^{(\kappa)}_j | g^{(\kappa)}_{j'} \rangle$ the overlap matrix between the GWPs.

To summarize, the dynamical equations (12.28)–(12.31) combine (i) a modified EOM for the $\boldsymbol{A}$ coefficients, taking into account the non-orthogonality of the hybrid configurations, (ii) the EOM for the primary-mode SPFs which is formally unchanged as compared with the standard MCTDH equations, and (iii) the new EOM for the GWP parameters, Eqs. (12.30)–(12.31).

The numerical scaling of the EOMs for the GWP parameters is determined by the solution of the coupled linear differential equations Eqs. (12.30)–(12.31), which involves an inversion of the $\boldsymbol{C}$ matrix in the standard implementation. Given the dimensionality $(nd)^2$ of the $\boldsymbol{C}$ matrix, where n is the number of GWPs in the κth subspace and d is the number of degrees of freedom combined in a given FG mode, the numerical effort scales with $(nd)^3$, growing rapidly with FG dimensionality.

12.3.1.3 vMCG Equations of Motion

In the vMCG scheme, pertaining to the wave function form Eq. (12.27), the EOMs Eq. (12.28) and Eq. (12.30)–(12.31) simplify – while keeping their formal appearance – and the mean-field structure disappears, along with the coupling to a primary subspace. The EOMs therefore take the following form, for the A coefficients (cf. Eq. (12.28)),

$$i\boldsymbol{S}\dot{\boldsymbol{A}} = (\boldsymbol{H} - i\boldsymbol{\tau})\boldsymbol{A} \tag{12.34}$$

and time-dependent GWP parameters $\boldsymbol{\Lambda}$ [19] (cf. Eq. (12.30)),

$$i\boldsymbol{C}\dot{\boldsymbol{\Lambda}} = \boldsymbol{Y} \tag{12.35}$$

where the matrix $\boldsymbol{C}$ and vector $\boldsymbol{Y}$ are now given as follows (cf. Eq. (12.31)),

$$\begin{aligned} C_{j\alpha,j'\beta} &= \rho_{jj'} \langle \partial_\alpha g_j | (\widehat{1} - \widehat{P}) | \partial_\beta g_{j'} \rangle \\ Y_{j\alpha} &= \sum_{j'} \langle \partial_\alpha g_j | (\widehat{1} - \widehat{P}) \widehat{H}_{jj'} | g_{j'} \rangle \end{aligned} \tag{12.36}$$

where $\rho_{jj'} = A^*_j A_{j'}$ and the projector $\widehat{P} = \sum_{jj'} |g_j\rangle ((\boldsymbol{S})^{-1})_{jj'} \langle g_{j'}|$ corresponds to Eq. (12.33) but now spans the full space. As compared with the general G-MCTDH scheme, the number of GWP subspaces is reduced to $f = 1$ and the primary modes have been eliminated (see also Eq. (12.27)). Due

to the cubic scaling of the Gaussian EOMs with $(nd)^3$, the increase of the vMCG numerical cost with the dimensionality of the system is a critical issue.

For further details of the vMCG approach and its use as a direct dynamics method with *on-the-fly* potentials, see Chapter 13.

12.3.2 2L-GMCTDH

Since parametrized functions such as GWPs are less flexible than the SPFs employed in standard MCTDH, a larger number of GWPs are required to correctly represent the wave function in the κth subspace. Furthermore, GWPs of FG type are not well suited to describe correlations between the combined degrees of freedom, which are represented as simple products of one-dimensional GWPs. (In contrast, the original concept of combined-mode TGs relies on the fact that the latter capture time-evolving correlations through the off-diagonal elements of the width matrix [19].) For these reasons, it is desirable to seek new wave function representations which combine the GWP representation with the flexibility of SPFs. This is the objective that led to the ML-GMCTDH – and specifically 2L-GMCTDH – method [25–27] which will be discussed in the following.

12.3.2.1 Wave Function Ansatz

As detailed in Ref. [25], the 2L-GMCTDH method aims to remedy the shortcomings of the factorizable FG Gaussian basis sets that are commonly employed in the G-MCTDH and vMCG approaches, by introducing a hierarchically structured wave function composed of two layers. The first ("outer") layer contains flexible SPFs, as in the conventional MCTDH wave function *ansatz*,

$$\Psi(\mathbf{x},t) = \sum_J A_J(t)\Phi_J(\mathbf{x},t) = \sum_J A_J(t) \prod_{\kappa=1}^{f} \chi_{j_\kappa}^{(\kappa)}(\mathbf{x}_\kappa,t) \tag{12.37}$$

Differently from MCTDH, the SPFs $\chi_{j_\kappa}^{(\kappa)}$ are now expressed as superpositions of multi-dimensional FG Gaussians, which form the second ("inner") layer,

$$\chi_j^{(\kappa)}(\mathbf{x}_\kappa,t) = \sum_L B_{j,L}^{(\kappa)}(t) G_L^{(\kappa)}(\mathbf{x}_\kappa,t) = \sum_L B_{j,L}^{(\kappa)}(t) \prod_{\mu=1}^{f_\kappa} g_{l_\mu}^{(\kappa,\mu)}(\mathbf{x}_{\kappa,\mu},t) \tag{12.38}$$

with multi-indices $L = (l_1, l_2, \ldots, l_{f^\kappa})$, $l_\mu \in \{1,2,\ldots,n_{\kappa,\mu}\}$, second-layer coefficients $B_{j,L}^{(\kappa)}(t)$ and multi-dimensional FGs $g_l^{(\kappa,\mu)}(\mathbf{x}_{\kappa,\mu},t)$. The latter are defined as in Eq. (12.25) except for the additional indices labeling each GWP in the (κ,μ) subspace. Hence, the wave function is of the same type as the 2L-MCTDH wave function of Eqs. (12.20)–(12.21), but now contains GWPs in the second layer.

In the above two-layer ansatz, the first-layer SPFs $\chi_j^{(\kappa)}$ clearly have more flexibility than conventional GWPs, due to the time-evolving $B_{j,L}^{(\kappa)}$ coefficients. Each second-layer GWP mode is uniquely specified by the index pair (κ,μ), where κ labels the first-layer mode and μ labels the second-layer mode. The index l_μ runs over the basis set of second-layer GWPs $g_{l_\mu}^{(\kappa,\mu)}$ which is of length $n_{\kappa,\mu}$.

In line with the above representation, two types of projector will appear in the equations of motion presented below: First, a projector onto the first-layer κth subspace:

$$\hat{P}^{(\kappa)} = \sum_j |\chi_j^{(\kappa)}\rangle\langle\chi_j^{(\kappa)}| \tag{12.39}$$

and, second, a projector onto the second-layer (κ,μ)th subspace:

$$\hat{P}^{(\kappa,\mu)} = \sum_{l,l'} |g_l^{(\kappa,\mu)}\rangle((\boldsymbol{S}^{(\kappa,\mu)})^{-1})_{ll'}\langle g_{l'}^{(\kappa,\mu)}| \tag{12.40}$$

where the latter is adapted to the non-orthogonal GWP representation, with $S_{ll'}^{(\kappa,\mu)} = \langle g_l^{(\kappa,\mu)} | g_{l'}^{(\kappa,\mu)} \rangle$ the overlap matrix elements between the GWPs.

Following conventional MCTDH notation, the wave function of Eqs. (12.37)–(12.38) can further be expressed in terms of products of SPFs and SHFs (see Eq. (12.11)),

$$\Psi = \sum_j \chi_j^{(\kappa)} \psi_j^{(\kappa)} = \sum_l g_l^{(\kappa,\mu)} \psi_l^{(\kappa,\mu)} \tag{12.41}$$

where $\psi_j^{(\kappa)}$ and $\psi_l^{(\kappa,\mu)}$ refer to first-layer and second-layer SHFs, respectively,

$$\begin{aligned} \psi_j^{(\kappa)} &= \sum_{J^{(\kappa)}} A_{J^{(\kappa:j)}} \Phi_{J^{(\kappa)}} \ , \\ \psi_l^{(\kappa,\mu)} &= \sum_j \left[\sum_{L^{(\mu)}} B_{j,L^{(\mu:l)}}^{(\kappa)} G_{L^{(\mu)}}^{(\kappa)} \right] \psi_j^{(\kappa)} \end{aligned} \tag{12.42}$$

Here we again used a shorthand notation for multi-indices and configurations, i.e., $J^{(\kappa)} = (j_1, \dots, j_{\kappa-1}, j_{\kappa+1}, \dots, j_f)$, $J^{(\kappa:j)} = (j_1, \dots, j_{\kappa-1}, j, j_{\kappa+1}, \dots, j_f)$ and $\Phi_{J^{(\kappa)}} = \prod_{\kappa' \neq \kappa} \chi_{j_{\kappa'}}^{(\kappa')}$. The quantities $L^{(\mu)}$, $L^{(\mu:l)}$ and $G_{L^{(\mu)}}^{(\kappa)}$ are defined analogously.

Using the above definitions, first- and second-layer density matrices are defined in terms of the overlap of the corresponding SHFs,

$$\rho_{jj'}^{(\kappa)} = \langle \psi_j^{(\kappa)} | \psi_{j'}^{(\kappa)} \rangle \quad , \quad \rho_{ll'}^{(\kappa,\mu)} = \langle \psi_l^{(\kappa,\mu)} | \psi_{l'}^{(\kappa,\mu)} \rangle \tag{12.43}$$

As is seen below, the SHFs also appear in the mean fields for both layers.

Importantly, the gauge freedom of MCTDH permits us to keep the first layer of the wave function representation orthogonal, while the second layer is composed of non-orthogonal GWP functions. As detailed in Ref. [33], the construction of Eqs. (12.37)–(12.38) can be continued to an arbitrary number of M SPF layers, the last of which is represented in terms of GWPs. This is entirely analogous to the standard multi-layer (ML-MCTDH) approach described in Section 12.2.3, except that the last layer is represented by non-orthogonal functions.

12.3.2.2 Equations of Motion

The equations of motion for the two-layer G-MCTDH ansatz of Eqs. (12.37)–(12.38) involve coupled non-linear differential equations for two types of time-dependent coefficients, i.e., first-layer coefficients $\boldsymbol{A}$ and second-layer coefficients $\boldsymbol{B}^{(\kappa)}$,

$$i\dot{\boldsymbol{A}} = \boldsymbol{H}\boldsymbol{A} \tag{12.44}$$

$$i\boldsymbol{S}^{(\kappa)}\dot{\boldsymbol{B}}^{(\kappa)} = [\boldsymbol{H}^{(\kappa)} - i\,\boldsymbol{\tau}^{(\kappa)}]\boldsymbol{B}^{(\kappa)} \tag{12.45}$$

along with the equations of motion for the second-layer GWP parameters $\boldsymbol{\Lambda}^{(\kappa,\mu)}$,

$$i\boldsymbol{C}^{(\kappa,\mu)}\dot{\boldsymbol{\Lambda}}^{(\kappa,\mu)} = \boldsymbol{Y}^{(\kappa,\mu)} \tag{12.46}$$

The dynamical equations Eqs. (12.44)–(12.46) combine (i) a standard MCTDH-like EOM for the first-layer $\boldsymbol{A}$ coefficients (due to the orthogonality of the first-layer SPFs), (ii) a G-MCTDH like equation for the second-layer $\boldsymbol{B}^{(\kappa)}$ coefficients (defined for a given κ-subspace), and (iii) a G-MCTDH like equation for the GWP parameters $\boldsymbol{\Lambda}^{(\kappa,\mu)}$, which are now defined for the (κ,μ)th subspace. By the two-layer structure, the GWP parameter evolution is restricted to low-dimensional (κ,μ)-subspaces, leading to a significant reduction of the numerical cost. This needs to be balanced against the cost increase due to the additional second-layer coefficients.

In further detail, in Eq. (12.44) for the first-layer coefficients, the Hamiltonian matrix comprises the elements

$$H_{JJ'} = \langle \Phi_J | \widehat{H} | \Phi_{J'} \rangle \tag{12.47}$$

i.e., Hamiltonian matrix elements in the basis of the first-layer configurations. Further, the following overlap matrix elements and Hamiltonian matrix elements pertaining to the κ-subspace appear in Eq. (12.45) for the second-layer coefficients,

$$S^{(\kappa)}_{jL,j'L'} = \delta_{jj'}\langle G^{(\kappa)}_L | G^{(\kappa)}_{L'}\rangle \,, \quad \tau^{(\kappa)}_{jL,j'L'} = \delta_{jj'}\langle G^{(\kappa)}_L | \partial_t G^{(\kappa)}_{L'}\rangle \tag{12.48}$$

$$H^{(\kappa)}_{jL,j'L'} = \left\langle G^{(\kappa)}_L \middle| (1-\widehat{P}^{(\kappa)})\left[(\boldsymbol{\rho}^{(\kappa)})^{-1}\widehat{\boldsymbol{H}}^{(\kappa)}\right]_{jj'} G^{(\kappa)}_{L'} \right\rangle \tag{12.49}$$

where $S^{(\kappa)}_{jL,j'L'}$ represents the overlap of second-layer configurations, and $\tau^{(\kappa)}_{jL,j'L'}$ represents the corresponding differential overlap element. In Eq. (12.49), $\rho^{(\kappa)}_{jj'}$ denotes the overlap of the first-layer SHFs, see Eq. (12.43), and $\widehat{H}^{(\kappa)}_{jj'} = \langle \psi^{(\kappa)}_j | \widehat{H} | \psi^{(\kappa)}_{j'}\rangle$ denotes the first-layer mean fields, while $\widehat{P}^{(\kappa)}$ is the orthogonal projector of Eq. (12.39).

Finally, in Eq. (12.46) for the GWP parameter vector, the $\boldsymbol{C}^{(\kappa,\mu)}$ matrix and $\boldsymbol{Y}^{(\kappa,\mu)}$ vector of the (κ,μ)-subspace have the following elements:

$$C^{(\kappa,\mu)}_{l\alpha,l'\beta} = \rho^{(\kappa,\mu)}_{ll'}\left\langle \partial_\alpha g^{(\kappa,\mu)}_l \middle| (1-\widehat{P}^{(\kappa,\mu)}) \middle| \partial_\beta g^{(\kappa,\mu)}_{l'} \right\rangle \tag{12.50}$$

$$Y^{(\kappa,\mu)}_{l\alpha} = \sum_{l'}\left\langle \partial_\alpha g^{(\kappa,\mu)}_l \middle| (1-\widehat{P}^{(\kappa,\mu)})\widehat{H}^{(\kappa,\mu)}_{ll'} \middle| g^{(\kappa,\mu)}_{l'} \right\rangle \tag{12.51}$$

where we again used a shorthand notation for the partial derivatives with respect to the GWP parameters, $\partial_\alpha g^{(\kappa,\mu)}_l = (\partial g^{(\kappa)}_l / \partial \Lambda^{(\kappa,\mu)}_{l\alpha})$. Further, $\rho^{(\kappa,\mu)}_{ll'}$ denotes the overlap of the second-layer single hole functions and $\widehat{H}^{(\kappa,\mu)}_{ll'} = \langle \psi^{(\kappa,\mu)}_l | \widehat{H} | \psi^{(\kappa,\mu)}_{l'}\rangle$ represents the second-layer mean fields.

Additional details regarding the derivation of Eqs. (12.44)–(12.46), along with the generalization to a multi-layer (ML-GMCTDH) method can be found in Ref. [25].

12.4 Quantum-Classical Multi-Configurational Approaches

The GWP parameter dynamics, which result from the application of the DFVP to a parametrized wave function [48], exhibit a symplectic structure – very similar to Hamilton's equations of classical mechanics. This structure follows from the analogy of the quantum Lagrangian $L = \langle H\rangle - i\sum_\alpha \langle \Psi | (\partial\Psi/\partial\Lambda_\alpha)\rangle \dot{\Lambda}_\alpha$ and the classical Lagrangian $L = H - p\dot{q}$ [45, 46, 48, 63]. In this sense, the Gaussian parameter dynamics is indeed pseudo-classical, and reduces to the classical equations of motion if applied to a single GWP [59, 60].

Given that the G-MCTDH approach can be interpreted as a hybrid quantum-*semiclassical* method, it is therefore natural to explore the quantum-*classical* limit of this approach. This line of investigation has been pursued in Refs. [33, 34] and will be summarized in the present section.

12.4.1 Quantum-Classical Limit of G-MCTDH

Specifically, we consider the following simple G-MCTDH type wave function, following Ref. [33],

$$\Psi(\boldsymbol{r},\mathbf{x},t) = \sum_{j=1}^{J}\sum_{l=1}^{L} B_{jl}(t)\varphi_j(\boldsymbol{r},t)g_l(\mathbf{x};\boldsymbol{\Lambda}_l(t)) \tag{12.52}$$

where $\boldsymbol{r}$ is the quantum variable while $\mathbf{x}$ is going to be taken to the classical limit. In the quantum-classical limit, the FG wave packet $g_l(\mathbf{x};\boldsymbol{\Lambda}_l(t))$ of Eq. (12.52) is effectively contracted to a phase-space point, by a scaling procedure detailed in Ref. [33], such that

$$\Psi^\varepsilon(\boldsymbol{r},\mathbf{x},t) = \sum_{j=1}^{J}\sum_{l=1}^{L} B_{jl}(t)\varphi_j(\boldsymbol{r},t)e^{\frac{i}{\varepsilon}S^{cl}_l(t)}g^\varepsilon_l(\mathbf{x};\boldsymbol{q}_l(t),\boldsymbol{p}_l(t)) \tag{12.53}$$

where $\varepsilon \to 0$ defines the classical limit [33], noting that we formally use $\hbar = 1$. The phase factors involve the classical action $S_l^{cl}(t)$, and the scaled FG wave packet is given as

$$g_l^\varepsilon(\mathbf{x}; \boldsymbol{q}_l(t), \boldsymbol{p}_l(t)) = N_\varepsilon \exp\left[(\mathbf{x} - \boldsymbol{q}_l(t))^T \boldsymbol{a}_l (\mathbf{x} - \boldsymbol{q}_l(t))/\varepsilon + \frac{i}{\varepsilon} \boldsymbol{p}_l^T(t)(\mathbf{x} - \boldsymbol{q}_l(t))\right] \tag{12.54}$$

In the following, we will consider the case of a quantum subspace comprising a discrete electronic basis $|n\rangle$ such that the ε-scaled wave function takes the form

$$|\Psi^\varepsilon(\mathbf{x}, t)\rangle = \sum_n^N \sum_l^L B_{nl}(t) e^{\frac{i}{\varepsilon} S_l^{cl}(t)} g_l^\varepsilon(\mathbf{x}; \boldsymbol{q}_l(t), \boldsymbol{p}_l(t))|n\rangle \tag{12.55}$$

In the context of MCTDH, this type of wave function *ansatz* corresponds to the so-called single-set form [3] for non-adiabatic dynamics. Eq. (12.55) also corresponds to a semiclassically scaled variant of a vMCG [23] wave function in single-set form [24]. Application of the DFVP now yields equations where the $B_{j,l}$ coefficients are found to be decoupled for different $l \neq l'$, i.e., configurations associated with different trajectories $(\boldsymbol{q}_l, \boldsymbol{p}_l)$ evolve independently:

$$i\dot{\boldsymbol{B}}_l = \boldsymbol{H}(\boldsymbol{q}_l)\, \boldsymbol{B}_l \tag{12.56}$$

where $\boldsymbol{B}_l = (B_{1l}, B_{2l}, \dots, B_{Nl})^T$ and the Hamiltonian matrix relates to the quantum subspace Hamiltonian augmented by the interaction potential evaluated at the classical position $\boldsymbol{q}_l$,

$$\widehat{H}_{nn'}(\boldsymbol{q}_l(t)) = h_{nn'} + \widehat{V}_{nn'}(\mathbf{x} = \boldsymbol{q}_l(t)) \tag{12.57}$$

where the Hamiltonian is assumed to take a generic diabatic form $\widehat{H}(\mathbf{x}) = \sum_{n,n'}(h_{nn'} + \widehat{V}_{nn'}(\mathbf{x}))|n\rangle\langle n'|$. From the DFVP, one further obtains classical evolution equations for the scaled Gaussians,

$$\dot{\boldsymbol{q}}_l = \boldsymbol{p}_l \qquad \dot{\boldsymbol{p}}_l = -\nabla_{\boldsymbol{q}_l} \bar{V}_l(\boldsymbol{q}_l) \tag{12.58}$$

where the mean-field potential is given as

$$\bar{V}_l(\boldsymbol{q}_l) = \left[\sum_n |B_{nl}|^2\right]^{-1} \sum_n \sum_{n'} B_{nl}^* B_{n'l} \langle n|\widehat{V}(\mathbf{x} = \boldsymbol{q}_l)|n'\rangle \tag{12.59}$$

In the following, we will employ the normalization condition $\sum_n |B_{nl}|^2 = 1$.

The above equations, i.e., Eq. (12.56) and Eqs. (12.58)–(12.59) correspond to the standard Ehrenfest equations [67, 68], such that the wave function of Eq. (12.55) can be interpreted as a sum of independently evolving Ehrenfest configurations $|\Phi_l^{qc}(\mathbf{x}, t)\rangle$,

$$|\Psi^\varepsilon(\mathbf{x}, t)\rangle = \sum_{l=1}^L |\Phi_l^{qc}(\mathbf{x}, t)\rangle \tag{12.60}$$

where the quantum-classical (qc) Ehrenfest configurations are given by

$$|\Phi_l^{qc}(\mathbf{x}, t)\rangle = \sum_{n=1}^N B_{nl}(t) e^{\frac{i}{\varepsilon} S_l^{cl}(t)} g_l^\varepsilon(\mathbf{x}; \boldsymbol{q}_l(t), \boldsymbol{p}_l(t))|n\rangle \tag{12.61}$$

Hence, the wave function Eq. (12.60) can be denoted a Multi-Configuration Ehrenfest (MCE) form (which is, however, different from the MCE approaches introduced in Refs. [31, 32]). In the quantum-classical wave function Eq. (12.60), phase information is kept by attaching the phase factor $\exp(\frac{i}{\varepsilon} S_l^{cl}(t))$ to each trajectory. However, due to the vanishing overlap of the trajectory-like g_l^ε functions, decoherence is effectively induced in the quantum subsystem.

12.4.2 Quantum-Classical Scheme with Finite-Width Wave Packets

In applications to molecular systems, the vibrational motions are not truly classical, but rather of semiclassical nature. Therefore, we return to the finite-width GWPs of Eq. (12.52), and combine these with the classical-limit dynamics of Eqs. (12.56)–(12.59). When doing so, it turns out, though, that the resulting dynamics is inconsistent since the normalization of the wave function is not conserved.

To impose norm conservation, we associate an additional set of coefficients A_l with each Ehrenfest configuration, and determine these coefficients from the DFVP. The modified wave function *ansatz* is given as follows [34],

$$|\Psi(\mathbf{x},t)\rangle = \sum_{l=1}^{L} A_l(t)|\Phi_l(\mathbf{x},t)\rangle \tag{12.62}$$

where the Ehrenfest configurations (now omitting the "qc" superscript) are given as

$$|\Phi_l(\mathbf{x},t)\rangle = \sum_{n=1}^{N} B_{nl}(t) g_l(\mathbf{x};\Lambda_l(t))\ |n\rangle \tag{12.63}$$

Here, the GWPs are taken to follow the classical equations of motion Eq. (12.58) even though they are in practice finite-width wave packets. All phase factors have now been absorbed into the wave function coefficients (as in Eq. (12.25)) such that the phase depending on the classical action no longer appears explicitly.

The resulting equations of motion read as follows [34],

$$i\boldsymbol{S}\dot{\boldsymbol{A}} = (\tilde{\boldsymbol{H}} - i\boldsymbol{\tau})\ \boldsymbol{A} \tag{12.64}$$

$$i\dot{\boldsymbol{B}}_l = \boldsymbol{H}(\boldsymbol{q}_l)\ \boldsymbol{B}_l \tag{12.65}$$

$$\dot{\boldsymbol{q}}_l = \boldsymbol{p}_l \qquad \dot{\boldsymbol{p}}_l = -\nabla_{\boldsymbol{q}_l}\bar{V}_l(\boldsymbol{q}_l) \tag{12.66}$$

where Eqs. (12.65)–(12.66) are unchanged from Eq. (12.56) and Eq. (12.58) while the new equation Eq. (12.64) determines the variational time evolution of the $\boldsymbol{A}$ coefficients.

All matrix elements appearing in the equation for the $\boldsymbol{A}$ coefficients are formulated in the basis of Ehrenfest configurations $|\Phi_l\rangle$ and read explicitly as follows,

$$\tilde{H}_{lk} = \langle\Phi_l|\hat{H}|\Phi_k\rangle = \sum_n \sum_{n'} B^*_{nl} B_{n'k} \langle n|\langle g_l|\hat{H}|g_k\rangle|n'\rangle \tag{12.67}$$

and similarly for the overlap matrix elements

$$S_{lk} = \langle\Phi_l|\Phi_k\rangle = \sum_n B^*_{nl} B_{nk} \langle g_l|g_k\rangle \tag{12.68}$$

and for the differential overlap matrix element,

$$\tau_{lk} = \langle\Phi_l|\dot{\Phi}_k\rangle = \sum_n (B^*_{nl}\dot{B}_{nk}\langle g_l|g_k\rangle + B^*_{nl}B_{nk}\langle g_l|\dot{g}_k\rangle) \tag{12.69}$$

In the last two equations, we used the orthogonality of the primitive electronic basis $|n\rangle$. In Eq. (12.69), the derivative matrix element on the right hand side can be evaluated using $\dot{g}_k = \sum_\alpha (\partial g_k/\partial\Lambda_{k\alpha})\dot{\Lambda}_{k\alpha}$ where the parameters $\Lambda_{k\alpha}$ constitute the parameter vector, $\boldsymbol{\Lambda}_k = \{\Lambda_{k\alpha}\}$. As a result, all matrix elements can be expressed as Gaussian moments.

Eqs. (12.65)–(12.66) are no longer fully variational, in contrast to Eq. (12.56) and Eq. (12.58) which are variational in the quantum-classical limit $\varepsilon \to 0$. However, the equation for the $\boldsymbol{B}$ coefficients is variational in a weaker sense: namely, the application of the DFVP to a single Ehrenfest configuration yields Eq. (12.56), if the phase relation Im $\dot{\mu} = \boldsymbol{p}\cdot\dot{\boldsymbol{q}}/\hbar$ is fulfilled.

12.4.3 Related Approaches

The wave function *ansatz* of Eqs. (12.62)–(12.63), with variational coefficients $\boldsymbol{A}_l$ but non-variational Ehrenfest configurations Φ_l is equivalent to the MCEv2 (i.e., MCE version 2) *ansatz* described by Shalashilin [32], except for the phase convention that is employed. This approach uses superpositions of independently evolving Ehrenfest configurations, in contrast to the MCEv1 approach [31] where these configurations are permitted to interact. While the latter approach is more flexible, MCEv2 is very convenient to use due to the independent evolution of the configurations. Notably, MCEv2 has been employed in *on-the-fly* AIMC-MCE calculations [69]. However, as detailed in Ref. [70] and also in our recent work [34], sampling strategies are a critical issue in the application of MCE.

A related hierarchy of Ehrenfest methods based on the vMCG ansatz has been presented in Ref. [73], while a recent approximate approach, starting from the exact factorization ansatz of Gross and co-workers [72] (see Chapter 17), can also relate vMCG to an Ehrenfest picture while retaining the GWP couplings [74]. In the context of *on-the-fly* calculations, Chapter 15 addresses the use of Ehrenfest methods.

In the general context of MCE type dynamics, it should be emphasized that energy is not a strictly conserved quantity, in contrast to fully variational schemes like G-MCTDH or vMCG. This shortcoming is common to all quantum methods that are based on classically moving Gaussians [34, 71].

12.5 How to use MCTDH & Co

The MCTDH method and its variants have been implemented in several publicly available packages, notably the Heidelberg MCTDH code [75] and more recently the QUANTICS package [76, 93]. The code of the latter package largely matches the Heidelberg MCTDH code, but also includes a GWP based branch with an implementation of the vMCG and G-MCTDH methods (where vMCG naturally arises as a special case of G-MCTDH as explained in Section 12.3.1). Both codes have been extensively optimized and employ, e.g., integration schemes – notably the constant mean field (CMF) integrator – that are specifically tailored to the MCTDH equations. Furthermore, various development versions exist; e.g., all 2L-GMCTDH and MCE calculations reported in this chapter have been carried out with an in-house code of the Frankfurt group.

All MCTDH type codes combine the integration of time-dependent wave function coefficients (e.g., Eq. (12.17)) with the integration of time-dependent SPFs (e.g., Eq. (12.18)), or the corresponding equations for GWP based approaches (i.e., Eq. (12.28) and Eq. (12.30) in the G-MCTDH case). In the MCTDH and ML-MCTDH methods, the SPFs are expressed in a primitive basis of time-independent DVR functions as mentioned in Section 12.2.2.2. In this sense, the standard MCTDH approach can be understood as a two-layered method. In contrast, the GWP based schemes obviate the distinction between the SPF basis and the primitive basis, since the GWPs are directly represented by the time-dependent parameter vector $\boldsymbol{\Lambda}(t)$, and the equations of motion are formulated in terms of the latter.

Initial conditions are specified in terms of one or more initially occupied configurations, whose coefficients and SPFs or GWPs are indicated. For example, in the case of FG type GWPs, initial positions and momenta have to be specified. Time propagation also requires initial conditions for initially unoccupied SPF or GWP functions, which are often fixed automatically by the program; e.g., in the case of GWPs, a regular distribution on a coordinate space or phase space grid could be chosen, with grid spacings that avoid initial linear dependencies.

Many useful details on the set-up of MCTDH calculations, which largely carry over to the related methods and implementations, can be found in the User's Guide of the Heidelberg MCTDH Package [77].

12.6 Synopsis and Application to Donor–Acceptor Complex

In this section, we illustrate the methods addressed above for a non-adiabatic coupling situation in a donor–acceptor (DA) system composed of an oligothiophene (OT_4) donor moiety and a fullerene (C_{60}) acceptor moiety. This system is a minimal model for the description of charge transfer in a paradigm DA system of organic photovoltaics, composed of poly-3-hexylthiophene (P3HT) and phenyl-C_{61} butyric acid methyl ester (PCBM) components [78–80]. The P3HT-PCBM system exhibits ultrafast charge transfer dynamics, on a time scale of $\sim$ 50 fs to 200 fs [79, 80]. In a previous study [41, 42] we employed a Linear Vibronic Coupling (LVC) Hamiltonian in conjunction with the MCTDH method to describe the ultrafast, coherent electron transfer dynamics in this system. The Hamiltonian was parametrized by electronic structure calculations using long-range-corrected density functional theory and the diabatization procedure described in Ref. [41], along with an *ab initio* generated spectral density of the phonon modes of the DA complex. In Refs. [81–84], extensions of this model to multiple electronic states are described.

12.6.1 Hamiltonian, Spectral Densities, and Potential Surfaces

The above mentioned LVC model describes the coupling of an excitonic donor state (OT_4^*-C_{60}), denoted XT, to a charge separated state (OT_4^+-C_{60}^-), denoted CT. As illustrated in Figure 12.3, this type of model can be employed either to molecular building blocks or else to coarse-grained units where, e.g., a fullerene "super-particle" is constructed [81–84].

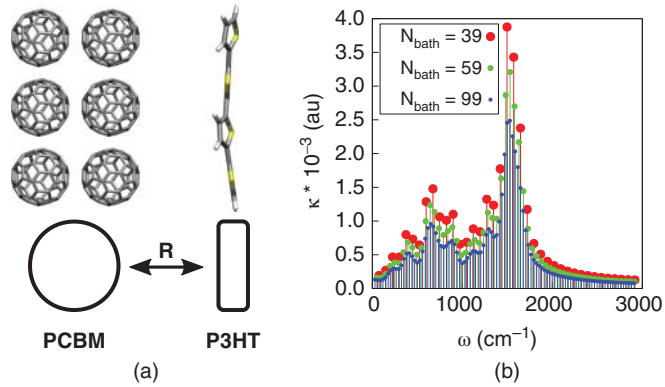


Figure 12.3 (a) Schematic representation of a OT_4-C_{60} donor–acceptor model complex representative of a P3HT-PCBM interface; here, the elementary building blocks are either molecular units or else coarse-grained "super-particles" as illustrated for the PCBM domain. (b) Vibronic couplings $\{\kappa_j\}$ pertaining to the OT_4-C_{60} system described in Refs. [41, 42], for various re-discretizations, for $N_{\text{bath}} = 39, 59, 99$ modes. The resulting discretized spectral densities $J(\omega) = (\pi/2)\sum_{j=1}^{N_{\text{bath}}} \kappa_j^2 \delta(\omega - \omega_j)$ yield the same dynamical evolution for times $t < \tau_P$ where $\tau_P = 2\pi/\Delta\omega$ is the Poincaré recurrence time.

The XT/CT two-state Hamiltonian is a function of N nuclear coordinates and can be cast in the form of a system-bath Hamiltonian [42],

$$\hat{H} = \hat{H}_0 + \hat{H}_R + \hat{H}_B \tag{12.70}$$

where $\hat{H}_0$ refers to the electronic part,

$$\hat{H}_0 = -\Delta_{\text{XT-CT}}|\text{CT}\rangle\langle\text{CT}| + \gamma(|\text{XT}\rangle\langle\text{CT}| + |\text{CT}\rangle\langle\text{XT}|) \tag{12.71}$$

while $\hat{H}_R$ is a vibronic part of the Hamiltonian depending on the inter-fragment distance coordinate R (see Figure 12.3(a)),

$$\hat{H}_R = \frac{\omega_R}{2}(\hat{R}^2 + \hat{P}^2) + \kappa_R\hat{R}|\text{CT}\rangle\langle\text{CT}| + \gamma_R\hat{R}(|\text{XT}\rangle\langle\text{CT}| + |\text{CT}\rangle\langle\text{XT}|) \tag{12.72}$$

and $\hat{H}_B$ represents a collection of $N_{\text{bath}} = N - 1$ intra-molecular "bath" modes pertaining to the oligothiophene and fullerene fragments. These are taken to be diagonally coupled to the CT state, assuming that the reference equilibrium geometry refers to the XT state,

$$\hat{H}_B = \sum_{i=1}^{N_{\text{bath}}} \frac{\omega_i}{2}(\hat{x}_i^2 + \hat{p}_i^2) + \sum_{i=1}^{N_{\text{bath}}} \kappa_i\hat{x}_i|\text{CT}\rangle\langle\text{CT}| \tag{12.73}$$

In Eqs. (12.70)–(12.73), $\Delta_{\text{XT-CT}}$ is the electronic offset, γ is a coordinate-independent diabatic coupling, κ_R and γ_R are electronically diagonal and off-diagonal vibronic couplings due to the intermolecular R-mode, and $\{\kappa_i\}$ describe electronically diagonal vibronic couplings of the bath modes. The above Hamiltonian employs mass- and frequency-weighted coordinates throughout. All parameters of the model are specified in Ref. [42].

As mentioned above, the vibronic coupling parameters κ_R and $\{\kappa_i\}$ express the displacement of the CT equilibrium geometry from the XT reference geometry. These couplings were determined by projecting the displacement between the XT and the CT minima onto normal mode coordinates for the separate OT_4^+ and C_{60}^- fragments, for 246 normal modes in total for both fragments [42]. The discrete distribution of electron–phonon couplings in the original normal-mode representation is subsequently used to construct a smooth spectral density function $J(\omega)$, representing a continuous density of modes that gives a better description of the high-dimensional polymer system. Here, we refer to a realization of this spectral density obtained with a Lorentzian broadening with width parameter $\Delta = 0.25\ \Delta_0$, where $\Delta_0 = 4.36\ 10^{-4}$ a.u. (96 cm^{-1}) corresponds to the RMS of the sampling distance of the original data (see Ref. [42] for details). Finally, the continuous spectral density $J(\omega)$ can be re-discretized with an arbitrary number of N_{bath} bath modes with an equidistant sampling interval $\Delta\omega$ such that several set-ups with variable dimensionality are generated [85, 86].

Using this re-discretization scheme, three realizations of $\hat{H}_B$ were obtained with $N_{\text{bath}} = 39$, 59 and 99 modes, shown in Figure 12.3(b). The corresponding Poincaré recurrence times $\tau_p = 2\pi/\Delta\omega$ are given as 446, 676, and 1135 fs. The different bath sizes allow us to explore the effect of the dimensionality on the performance and the convergence of the various propagation methods.

In Figure 12.4, the diabatic and adiabatic potential energy surfaces (PES) pertaining to the LVC Hamiltonian are plotted as a function of the inter-fragment coordinate $\hat{R}$ and an effective "Brownian oscillator" mode defined as $\hat{Q} = (1/D)\sum_i \kappa_i\hat{x}_i$ [87] which subsumes the coupling of the bath modes to the electronic subsystem and determines the short-time dynamics. In our dynamical simulations, the wave packet starts at the minimum of the XT potential, representing a relaxed exciton state. Due to the fact that the XT-CT coupling and the electronic offset take similar values, $\gamma \simeq \Delta_{\text{XT-CT}}$, the wave packet starts to oscillate between the two diabatic states. Eventually, the system relaxes to a quasi-stationary state which is predominantly of CT character (~90%) with

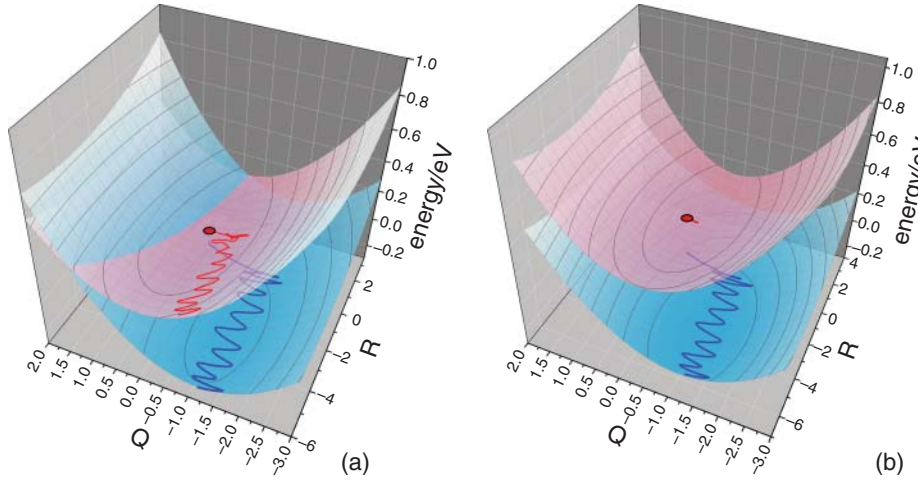


Figure 12.4 Diabatic (a) and adiabatic (b) representations of the PESs pertaining to the Hamiltonian Eqs. (12.70)–(12.73) shown as a function of the R coordinate and of an effective mode defined as $\hat{Q} = 1/D\sum_i \kappa_i \hat{x}_i$ [87] which subsumes the coupling of the bath modes to the electronic subsystem. The trajectories correspond to state-specific expectation values in the diabatic versus adiabatic representation, departing from the initial condition on the XT surface, marked by a circle. The concerted oscillatory motion in the diabatic representation (panel (a)) indicates that the wave packet exhibits a sustained coherent superposition of the portions belonging to the two diabatic states. In contrast, the residual wave packet portion in the upper adiabatic state essentially remains stationary (see panel (b)). Adiabatic coordinate expectation values were calculated for a 20-mode system using a simplified diabatic-to-adiabatic transformation along the paths defined by the full set of time-evolving coordinate expectation values. Reprinted from Ref. [34] with the permission of AIP Publishing.

a non-negligible XT admixture (∼10%). When evolving towards the quasi-stationary state, about 0.5 eV of excess energy is transferred to the bath modes. At the same time, the R coordinate is displaced to negative values, indicating a reduction of the inter-fragment distance. As illustrated by the time-evolving state-specific expectation values in Figure 12.4, concerted oscillations of the XT and CT portions of the wave packet are observed throughout the simulation interval.

12.6.2 Ultrafast Coherent Charge Transfer Dynamics

In Figure 12.5, time-evolving populations of the initially occupied XT state are shown along with the concomitant dynamics of the electronic coherences,

$$\rho_{\mathrm{XT,CT}}(t) = \mathrm{Tr}\{|\mathrm{CT}\rangle\langle\mathrm{XT}|\hat{\rho}(t)\} \tag{12.74}$$

Here, the imaginary part $\mathrm{Im}\rho_{\mathrm{XT,CT}}$ determines the transient state-to-state population flux, $\Gamma_{\mathrm{XT,CT}} = (-2\gamma/\hbar)\mathrm{Im}\rho_{\mathrm{XT,CT}}$ where γ is the diabatic coupling, while the real part $\mathrm{Re}\rho_{\mathrm{XT,CT}}$ captures the quasi-stationary XT-CT superposition that was mentioned above [42].

As illustrated in Figure 12.5, the coherent evolution between the two states plays an important role throughout the dynamics. The initial oscillatory decay of the XT state mirrors the profile of $\mathrm{Im}\rho_{\mathrm{XT,CT}}(\mathrm{t})$ which decays to zero within about 50 fs. Beyond this time, the XT state population correlates with the real part of the electronic coherence, $\mathrm{Re}\rho_{\mathrm{XT,CT}}$ (panel (b)), which tends towards a quasi-stationary value indicating that the lower adiabatic state is reached within approximately 100 fs.

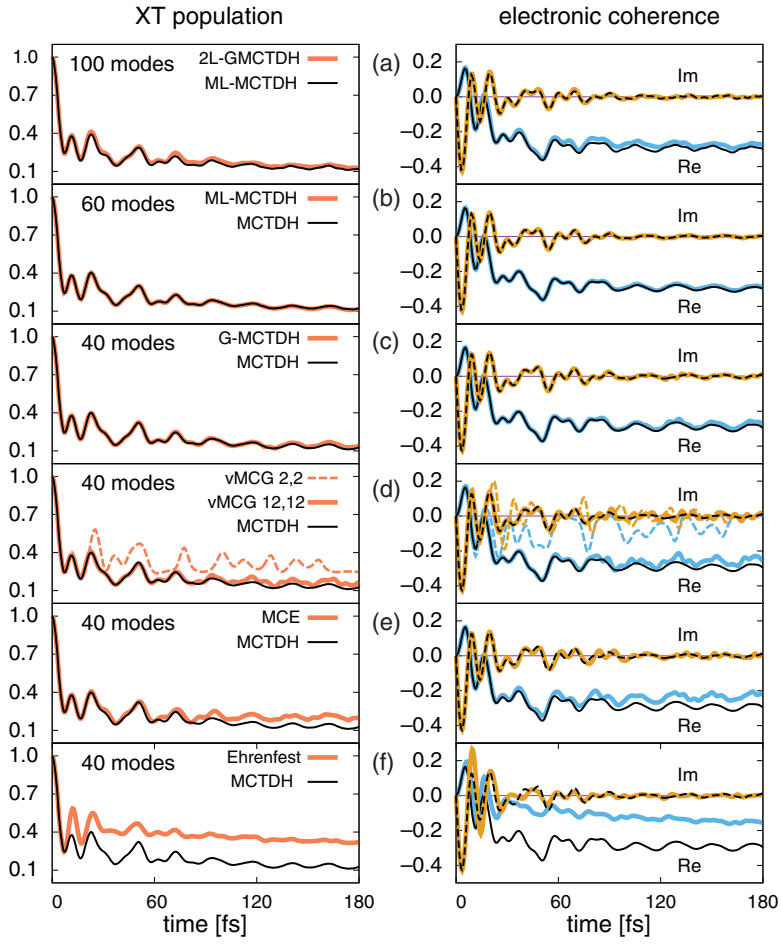


Figure 12.5 Time-dependent XT state population (left panels) and real and imaginary parts of the electronic coherence $\rho_{XT,CT}$ (right panels) are shown for different sizes of the bath ($N = N_{bath}+1 = 100$, 60, 40 modes) and different propagation methods, from MCTDH and ML-MCTDH to G-MCTDH and 2L-GMCTDH, MCE and statistical Ehrenfest. A detailed discussion of the results is given in the text.

In the following, we comment on the performance of the various methods for this system.

12.6.3 Comparison of Methods

The results shown in Figure 12.5 range from a 100-mode realization to 60-mode and 40-mode realizations of the system. As mentioned above, the resulting dynamics is identical *a priori*, since the observation window lies below the respective Poincaré times.

In panel (a), for the 100-mode system, a converged 2L-GMCTDH calculation is compared with a ML-MCTDH reference propagation comprising M=6 layers. (Details of the latter calculation are provided in the Supp. Mat. of Ref. [27]). Both for the ML-MCTDH reference calculation and for the 2L-GMCTDH calculation, a single-set type wave function is employed. The partitioning of the 2L-GMCTDH calculation is illustrated in Figure 12.6; here, the bath modes are combined into 20- to 30-dimensional first-layer particles, which are in turn split into three or four second-layer particles. The intermolecular R mode is taken separately as a one-dimensional particle.

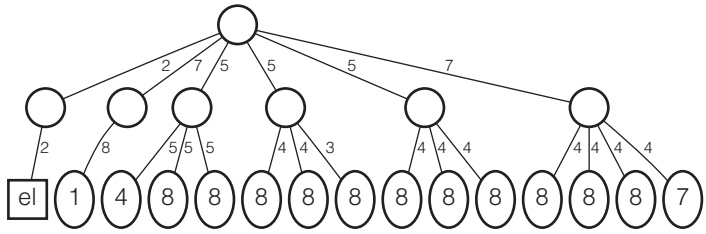


Figure 12.6 Two-layer tree for the 2L-GMCTDH calculation for a 100-mode system using a single-set set-up [27]. The numbers indicated inside the ellipses representing second-layer particles give the number of combined modes, and the numbers shown at the edges of the graphs indicate the number of first-layer and second-layer SPFs. Reprinted from Ref. [27] with the permission of AIP Publishing.

For the 60-mode system, a conventional MCTDH calculation is feasible, which compares well with the ML-MCTDH result, as shown in panel (b). However, a much higher computational cost arises for the MCTDH calculation (see Table 12.1).

In the above calculations for a rather large number of degrees of freedom, convergence is contingent upon the multi-configurational partitioning of the wave function. For the more approximate methods that do not employ any partitioning into subspaces – notably the vMCG and MCE approaches – convergence is out of reach for 60 or more modes. Hence, results for these more approximate methods are now discussed for a 40-mode realization, where these methods are able to approach the exact result.

Table 12.1 Complementary to the discussion of Section 12.6.3, the performance of different methods in terms of CPU times and memory requirements is illustrated. The Heidelberg MCTDH code [75] was employed for MCTDH and ML-MCTDH calculations, along with a predecessor version of the QUANTICS package [76] for the G-MCTDH and vMCG calculations, as well as an in-house code (Frankfurt University) for the 2L-GMCTDH and MCE calculations.

Number of modes	Method	Number of coefficients	CPU time [s]	RAM [MB]
40	MCTDH	143 616	1310	264
40	G-MCTDH hybrid	143 616	791	34
40	G-MCTDH GWP only	143 616	4966	35
40	G-MCTDH GWP only	2 890 080	37 167	652
40	vMCG (2,2)	4	3879	2
40	vMCG (6,6)	12	33 455	14
40	vMCG (10,10)	20	109 161	37
40	vMCG (12,12)	24	189 871	53
40	vMCG (14,14)	28	311 647	72
40	MCE, 2000 traj.	2000	607 860	57
60	MCTDH	18 040 320	56 989	3394
60	ML-MCTDH, M=6	65 433	9152	13
100	2L-GMCTDH, single-set	15 021	14 588	20
100	ML-MCTDH, M=6	46 501	36 488	13

Panels (c)–(f) show various results using the MCTDH, G-MCTDH, vMCG, MCE, and Ehrenfest approaches. Among the GWP-based methods, it turns out that only the G-MCTDH calculations are well converged (panel (c)). The vMCG results are rather poor for two GWPs per state, but improve quite significantly for 12 GWPs per state (panel (d)). Interestingly, the vMCG calculation with two GWPs shows a large deviation of the real part of the electronic coherence, $\mathrm{Re}\rho_{\mathrm{XT,CT}}$, and is apparently not suitable to describe the XT-CT quasi-stationary coherent superposition that characterizes the asymptotic state.

The MCE and Ehrenfest calculations shown in panels (e)–(f) illustrate that approximate convergence is feasible for MCE, using 2000 GWPs, with an enhanced sampling approach as detailed in Ref. [34]. In contrast, the standard statistical Ehrenfest approach – here employed for 5000 trajectories, using Wigner function sampling – is not able to correctly describe the system's time evolution. The reason for this is clearly the failure of the Ehrenfest method to correctly describe the XT-CT coherent superposition state, as is seen from the incorrect representation of $\mathrm{Re}\rho_{\mathrm{XT,CT}}$ on the right hand side of panel (f). This is not unexpected, given the failure of the Ehrenfest method to capture detailed balance [88, 89]. On the shortest time scales, the deviations are even larger than in the vMCG calculations restricted to two GWPs. A detailed assessment of the MCE vs. Ehrenfest schemes for the 40-mode case is given in Ref. [34].

Table 12.1 summarizes the performance of the different methods, in terms of CPU times and memory requirements. All GWP based methods are favorable in terms of memory requirements, but CPU times vary significantly depending on the degree of mode combination and the hierarchical construction of the wave function. Clearly, multi-layer methods – both ML-MCTDH and 2L-GMCTDH – show a very favorable scaling.

12.7 Conclusions and Outlook

Over the past few decades, high-dimensional quantum dynamics has made big strides due to the advent of the MCTDH method. The variational, multi-configurational approach is a systematic route towards accurate, correlated quantum dynamics. Recently introduced multi-layer variants permit the treatment of correlations to be carried out for hundreds of modes. Current progress at the interface of chemistry, physics, and mathematics, benefits from the connections between MCTDH and low-rank tensor approximation schemes, leading to new developments.

While high-dimensional wave function methods are often seen as disjoint from trajectory-based methods, this chapter attempts to show that the multi-configurational approach permits a smooth transition from a fully quantum, to a semiclassical and quantum-classical treatment. This transition is best understood from the viewpoint of the Gaussian-based MCTDH method, i.e., G-MCTDH, whose EOMs involve non-classically coupled evolution of the moving Gaussian basis. This chapter shows how several variants – including vMCG and ML-GMCTDH – emerge and how these methods perform as a function of the dimensionality and correlations of the system. Furthermore, an additional step is made towards classically evolving Gaussians, as employed in the multi-configuration Ehrenfest scheme [34], and we analyzed to what extent these different types of methods are able to capture quantum coherent dynamics.

Regarding the spin-boson type system that was chosen as an application to non-adiabatic dynamics in many dimensions, we conclude that despite the near-harmonic dynamics in the bath subspaces, the system provides an informative testbed for a comparative study. Besides the necessity of employing multi-configurational approaches to achieve convergence in the higher-dimensional simulations, a key criterion for the convergence of the dynamics is the correct representation of the

electronic coherence evolution. While the statistical Ehrenfest approach clearly falls short of this requirement, the MCE method is able to reach convergence, albeit at a high numerical cost.

Due to the multi-configurational wave function form employed in MCTDH type methods, potential energy surfaces need to be fitted to a sum-over-products form. While this is a restriction *a priori* – especially in the context of *on-the-fly* dynamics – various novel PES fitting strategies, including neural network fitting [90, 91] and other schemes, potentially alleviate this constraint. Hence, the direct use of MCTDH and related schemes in *on-the-fly* applications is in reach [92, 94]. In this context, complementary approaches can be pursued where trajectory-based *on-the-fly* information can be employed as input to accurate quantum propagation.

Acknowledgments

We thank Pierre Eisenbrandt, Matthias Ruckenbauer, Sarah Römer, David Picconi, Tianji Ma, and Robert Binder for their contributions to the work reported in this chapter. The German–Israeli Foundation for Scientific Research and Development is acknowledged for support of this work under grant number GIF I-1337-302.5/2016. M.B. gratefully acknowledges fellowship support by the Alexander von Humboldt Foundation.

References

1 Meyer, H.-D., Manthe, U., and Cederbaum, L.S. (1990). The multiconfigurational time-dependent Hartree approach. *Chem. Phys. Lett.* 165: 73.

2 Manthe, U., Meyer, H.-D., and Cederbaum, L.S. (1992). Wave-packet dynamics within the multiconfiguration Hartree framework: general aspects and application to NOCl. *J. Chem. Phys.* 97: 3199.

3 Beck, M.H., Jäckle, A., Worth, G.A., and Meyer, H.-D. (2000). The multiconfiguration time-dependent Hartree (MCTDH) method: a highly efficient algorithm for propagating wave packets. *Phys. Rep.* 324: 1.

4 Meyer, H.D., Gatti, F., and Worth, G.A. (2009). *Multidimensional Quantum Dynamics: MCTDH Theory and Applications*. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA.

5 Gatti, F., Lasorne, B., Meyer, H., and Nauts, A. (2017). *Applications of Quantum Dynamics in Chemistry, Lecture Notes in Chemistry 98*. Switzerland: Springer.

6 Worth, G.A., Meyer, H., Köppel, H. et al. (2008). Using the MCTDH wave packet propagation method to describe multimode non-adiabatic dynamics. *Int. Rev. Phys. Chem.* 27: 569.

7 Kato, T. and Kono, H. (2004). Time-dependent multiconfiguration theory for electronic dynamics of molecules in an intense laser field. *Chem. Phys. Lett.* 392: 533.

8 Zanghellini, J., Kitzler, M., Brabec, T., and Scrinzi, A. (2004). Testing the multiconfiguration time-dependent Hartree-Fock method. *J. Phys. B Atomic Mol. Phys.* 37: 763.

9 Nest, M., Padmanaban, R., and Saalfrank, P. (2007). Time-dependent approach to electronically excited states of molecules with the multiconfiguration time-dependent Hartree-Fock method. *J. Chem. Phys.* 126: 214106.

10 Alon, O.E., Streltsov, A.E., and Cederbaum, L.S. (2008). Multiconfigurational time-dependent Hartree method for bosons: many-body dynamics of bosonic systems. *Phys. Rev. A* 77: 033613.

11 Wang, H. and Thoss, M. (2003). Multilayer formulation of the multiconfiguration time-dependent Hartree theory. *J. Chem. Phys.* 119: 1289.

12 Wang, H. (2015). Multilayer multiconfiguration time-dependent Hartree theory. *J. Phys. Chem. A* 119: 7951.

13 Manthe, U. (2008). A multilayer multiconfigurational time-dependent Hartree approach for quantum dynamics on general potential energy surfaces. *J. Chem. Phys.* 128: 164116.

14 Vendrell, O. and Meyer, H.-D. (2011). Multilayer multiconfiguration time-dependent Hartree method: implementation and applications to a Henon-Heiles Hamiltonian and to pyrazine. *J. Chem. Phys.* 134: 044135.

15 Raab, A., Burghardt, I., and Meyer, H. (1999). The multiconfiguration time-dependent Hartree method generalized to the propagation of density operators. *J. Chem. Phys.* 111: 8759.

16 Raab, A. and Meyer, H.-D. (2000). A numerical study on the performance of the multiconfiguration time-dependent Hartree method for density operators. *J. Chem. Phys.* 112: 10718.

17 Heller, E.J. (1975). Time-dependent approach to semiclassical dynamics. *J. Chem. Phys.* 62: 1544.

18 Heller, E.J. (2018). *The Semiclassical Way to Dynamics and Spectroscopy*. Princeton University Press.

19 Burghardt, I., Meyer, H.-D., and Cederbaum, L.S. (1999). Approaches to the approximate treatment of complex molecular systems by the multiconfiguration time-dependent Hartree method. *J. Chem. Phys.* 111: 2927.

20 Burghardt, I., Nest, M., and Worth, G.A. (2003). Multiconfigurational system-bath dynamics using Gaussian wave packets: energy relaxation and decoherence induced by a finite-dimensional bath. *J. Chem. Phys.* 119: 5364.

21 Burghardt, I., Giri, K., and Worth, G.A. (2008). Multimode quantum dynamics using Gaussian wave packets: the Gaussian-based multiconfiguration time-dependent Hartree (G-MCTDH) method applied to the absorption spectrum of pyrazine. *J. Chem. Phys.* 129: 174104.

22 Worth, G.A. and Burghardt, I. (2003). Full quantum mechanical molecular dynamics using Gaussian wave packets. *Chem. Phys. Lett.* 368: 502.

23 Worth, G.A., Robb, M.A., and Burghardt, I. (2004). A novel algorithm for non-adiabatic direct dynamics using variational Gaussian wave packets. *Faraday Discuss.* 127: 307.

24 Richings, G.W., Polyak, I., Spinlove, K.E. et al. (2015). Quantum dynamics simulations using Gaussian wave packets: the vMCG method. *Int. Rev. Phys. Chem.* 34: 269.

25 Römer, S., Ruckenbauer, M., and Burghardt, I. (2013). Gaussian-based multiconfiguration time-dependent Hartree: a two-layer approach. I. Theory. *J. Chem. Phys.* 138: 064106.

26 Eisenbrandt, P., Ruckenbauer, M., Römer, S., and Burghardt, I. (2018). Gaussian-based multiconfiguration time-dependent Hartree: a two-layer approach. II. Application to vibrational energy transport in a molecular chain. *J. Chem. Phys.* 149: 174101.

27 Eisenbrandt, P., Ruckenbauer, M., and Burghardt, I. (2018). Gaussian-based multiconfiguration time-dependent Hartree: a two-layer approach. III. Application to non-adiabatic dynamics in a charge transfer complex. *J. Chem. Phys.* 149: 174102.

28 Heller, E.J. (1981). The semiclassical way to molecular spectroscopy. *Acc. Chem. Res.* 14: 368.

29 Heller, E.J. (1981). Frozen Gaussians: a very simple semiclassical approximation. *J. Chem. Phys.* 75: 2923.

30 Virshup, A.M., Punwong, C., Pogorelov, T.V. et al. (2009). Photodynamics in complex environments: Ab Initio multiple spawning quantum mechanical/molecular mechanical dynamics. *J. Phys. Chem. B* 113: 3280.

31 Shalashilin, D.V. (2009). Quantum mechanics with the basis set guided by Ehrenfest trajectories: theory and application to spin-boson model. *J. Chem. Phys.* 130: 244101.

32 Shalashilin, D.V. (2010). Non-adiabatic dynamics with the help of multiconfigurational Ehrenfest method: improved theory and fully quantum 24D simulation of pyrazine. *J. Chem. Phys.* 132: 244111.

33 Römer, S. and Burghardt, I. (2013). Towards a variational formulation of mixed quantum-classical molecular dynamics. *Mol. Phys.* 111: 3618.

34 Ma, T., Bonfanti, M., Eisenbrandt, P. et al. (2018). Multi-configurational Ehrenfest simulations of ultrafast non-adiabatic dynamics in a charge-transfer complex. *J. Chem. Phys.* 149: 244107.

35 Hackbusch, W. and Kühn, S. (2009). A new scheme for the tensor representation. *J. Fourier Anal. Appl.* 15: 706.

36 Lubich, C., Rohwedder, T., Schneider, R., and Vandereycken, B. (2013). Dynamical approximation by hierarchical Tucker and tensor-train tensors. *SIAM J. Matrix Anal. Appl.* 34: 470.

37 Lubich, C. (2008). *From Quantum to Classical Molecular Dynamics: Reduced Models and Numerical Analysis*. Zuerich, Switzerland: European Mathematical Society Publishing House.

38 Lubich, C. (2014). Time integration in the multiconfiguration time-dependent Hartree method of molecular quantum dynamics. *App. Math. Res. Express* 2015: 311.

39 Kloss, B., Burghardt, I., and Lubich, C. (2017). Implementation of a novel projector-splitting integrator for the multiconfigurational time-dependent Hartree approach. *J. Chem. Phys.* 146: 174107.

40 Bonfanti, M. and Burghardt, I. (2018). Tangent space formulation of the Multi-configuration Time-dependent Hartree equations of motion: the projector-splitting algorithm revisited. *Chem. Phys.* 515: 252.

41 Tamura, H., Burghardt, I., and Tsukada, M. (2011). Exciton dissociation at thiophene/fullerene interfaces: the electronic structures and quantum dynamics. *J. Phys. Chem.* 115: 10205.

42 Tamura, H., Martinazzo, R., Ruckenbauer, M., and Burghardt, I. (2012). Quantum dynamics of ultrafast charge transfer at an oligothiophene-fullerene heterojunction. *J. Chem. Phys.* 137: 22A540.

43 Dirac, P.A.M. (1930). Note on exchange phenomena in the Thomas atom. *Math. Proc. Camb. Philos. Soc.* 26: 376.

44 Frenkel, J. (1932). *Wave Mechanics: Elementary Theory*. Clarendon Press Oxford.

45 Rowe, D.J., Ryman, A., and Rosensteel, G. (1980). Many-body quantum mechanics as a symplectic dynamical system. *Phys. Rev. A* 22: 2362.

46 Broeckhove, J., Lathouwers, L., Kesteloot, E., and Van Leuven, P. (1988). On the equivalence of time-dependent variational principles. *Chem. Phys. Lett.* 149: 547.

47 McLachlan, A. (1964). A variational solution of the time-dependent Schrödinger equation. *Mol. Phys.* 8: 39.

48 Kramer, P. and Saraceno, M. (1981). *Geometry of the Time-Dependent Variational Principle in Quantum Mechanics. Lecture Notes in Physics*, Vol. 140. Berlin Springer Verlag.

49 Worth, G.A., Meyer, H.-D., and Cederbaum, L.S. (1998). Relaxation of a system with a conical intersection coupled to a bath: a benchmark 24-dimensional wave packet study treating the environment explicitly. *J. Chem. Phys.* 109: 3518.

50 Bachmayr, M., Schneider, R., and Uschmajew, A. (2016). Tensor networks and hierarchical tensors for the solution of high-dimensional partial differential equations. *Found. Comput. Math.* 16: 1423.

51 Manthe, U. (2015). The multiconfigurational time-dependent Hartree approach revisited. *J. Chem. Phys.* 142: 244109.

52 Meyer, H.-D. and Wang, H. (2018). On regularizing the MCTDH equations of motion. *J. Chem. Phys.* 148: 124105.

53 Raab, A., Worth, G.A., Meyer, H.-D., and Cederbaum, L.S. (1999). Molecular dynamics of pyrazine after excitation to the S_2 electronic state using a realistic 24-mode model Hamiltonian. *J. Chem. Phys.* 110: 936.

54 Worth, G., Robb, M., and Lasorne, B. (2008). Solving the time-dependent Schrödinger equation for nuclear motion in one step: direct dynamics of non-adiabatic systems. *Mol. Phys.* 106: 2077.

55 Picconi, D., Cina, J.A., and Burghardt, I. (2019). Quantum dynamics and spectroscopy of dihalogens in solid matrices. I. Efficient simulation of the photodynamics of the embedded I_2Kr_{18} cluster using the G-MCTDH method. *J. Chem. Phys.* 150: 064111.

56 Picconi, D., Cina, J.A., and Burghardt, I. (2019). Quantum dynamics and spectroscopy of dihalogens in solid matrices. II. Theoretical aspects and G-MCTDH simulations of time-resolved coherent Raman spectra of Schrödinger cat states of the embedded I_2Kr_{18} cluster. *J. Chem. Phys.* 150: 064112.

57 Picconi, D. and Burghardt, I. (2019). Open system dynamics using Gaussian-based multiconfigurational time-dependent Hartree wave functions: application to environment-modulated tunneling. *J. Chem. Phys.* 150: 224106.

58 Lee, S. and Heller, E.J. (1982). Exact time-dependent wave packet propagation: application to the photodissociation of methyl iodide. *J. Chem. Phys.* 76: 3035.

59 Sawada, S., Heather, R., Jackson, B., and Metiu, H. (1985). A strategy for time dependent quantum mechanical calculations using a Gaussian wave packet representation of the wave function. *J. Chem. Phys.* 83: 3009.

60 Sawada, S. and Metiu, H. (1986). A multiple trajectory theory for curve crossing problems obtained by using a Gaussian wave packet representation of the nuclear motion. *J. Chem. Phys.* 84: 227.

61 Shalashilin, D.V. and Child, M.S. (2001). Description of tunneling with the help of coupled frozen Gaussians. *J. Chem. Phys.* 114: 9296.

62 Shalashilin, D.V. and Child, M.S. (2004). The phase space CCS approach to quantum and semiclassical molecular dynamics for high-dimensional systems. *Chem. Phys.* 304: 103.

63 Shalashilin, D.V. and Burghardt, I. (2008). Gaussian-based techniques for quantum propagation from the time-dependent variational principle: formulation in terms of trajectories of coupled classical and quantum variables. *J. Chem. Phys.* 129: 084104.

64 Klauder, J.R. and Skagerstam, B.S. (1985). *Coherent States: Applications in Physics and Mathematical Physics*. World Scientific.

65 Mendive-Tapia, D., Lasorne, B., Worth, G.A. et al. (2012). Towards converging non-adiabatic direct dynamics calculations using frozen-width variational Gaussian product basis functions. *J. Chem. Phys.* 137: 22A548.

66 Bonfanti, M., Petersen, J., Eisenbrandt, P. et al. (2018). Computation of the $S_1 \leftarrow S_0$ vibronic absorption spectrum of formaldehyde by variational Gaussian wave packet and semiclassical IVR methods. *J. Chem. Theory Comput.* 14: 5310.

67 Delos, J.B., Thorson, W.R., and Knudson, S.K. (1972). Semiclassical theory of inelastic collisions. I. Classical picture and semiclassical formulation. *Phys. Rev. A* 6: 709.

68 Billing, G.D. (1983). On the use of Ehrenfest's theorem in molecular scattering. *Chem. Phys. Lett.* 100: 535.

69 Makhov, D.V., Symonds, C., Fernandez-Alberti, S., and Shalashilin, D.V. (2017). Ab initio quantum direct dynamics simulations of ultrafast photochemistry with multiconfigurational Ehrenfest approach. *Chem. Phys.* 493: 200.

70 Symonds, C., Kattirtzi, J.A., and Shalashilin, D.V. (2018). The effect of sampling techniques used in the multiconfigurational Ehrenfest method. *J. Chem. Phys.* 148: 184113.
71 Habershon, S. (2012). Linear dependence and energy conservation in Gaussian wave packet basis sets. *J. Chem. Phys.* 136: 014109.
72 Abedi, A., Maitra, N.T., and Gross, E. (2010). Exact factorization of the time-dependent electron-nuclear wave function. *Phys. Rev. Lett.* 105: 123002.
73 Spinlove, K.E., Vacher, M., Bearpark, M. et al. (2017). Using quantum dynamics simulations to follow the competition between charge migration and charge transfer in polyatomic molecules. *Chem. Phys.* 482: 52.
74 Jenkins, A.J., Spinlove, K.E., Vacher, M. et al. (2018). The Ehrenfest method with fully quantum nuclear motion (Qu-Eh): application to charge migration in radical cations. *J. Chem. Phys.* 149: 094108.
75 Worth, G.A., Beck, M.H., Jäckle, A., and Meyer, H.-D. (2016). The MCTDH Package, Version 8.5.5. See http://mctdh.uni-hd.de.
76 Worth, G.A., Giri, K., Richings, G.W. et al. (2015). The QUANTICS Package, Version 1.1. Birmingham, UK: University of Birmingham.
77 Worth, G.A, Beck, M.H., Jäckle, A., and Meyer, H.-D. (2018). The Heidelberg MCTDH Package: a set of programs for multidimensional quantum dynamics – User's Guide, version 8, release 4, revision 16.
78 Sariciftci, N.S., Smilowitz, L., Heeger, A.J., and Wudl, F. (1992). Photoinduced electron transfer from a conducting polymer to buckminsterfullerene. *Science* 258: 1474.
79 Brabec, C.J., Zerza, G., Cerullo, G. et al. (2001). Tracing photoinduced electron transfer process in conjugated polymer/fullerene bulk heterojunctions in real time. *Chem. Phys. Lett.* 340: 232.
80 Pensack, R.D. and Asbury, J.B. (2010). Beyond the adiabatic limit: charge photogeneration in organic photovoltaic materials. *J. Phys. Chem. Lett.* 1: 2255.
81 Tamura, H. and Burghardt, I. (2013). Ultrafast charge separation in organic photovoltaics enhanced by charge delocalization and vibronically hot exciton dissociation. *J. Am. Chem. Soc.* 135: 16364.
82 Huix-Rotllant, M., Tamura, H., and Burghardt, I. (2015). Concurrent effects of delocalization and internal conversion tune charge separation at donor–acceptor polymer heterojunctions. *J. Phys. Chem. Lett.* 6: 1702.
83 Polkehn, M., Tamura, H., and Burghardt, I. (2018). Impact of charge transfer excitons in regioregular polythiophene on the charge separation at polythiophene-fullerene heterojunctions. *J. Phys. B Atomic Mol. Opt. Phys.* 51: 014003.
84 Polkehn, M., Eisenbrandt, P., Tamura, H., and Burghardt, I. (2018). Quantum dynamical studies of ultrafast charge separation in nanostructured organic polymer materials: effects of vibronic interactions and molecular packing. *Int. J. Quantum Chem.* 118: e25502.
85 Martinazzo, R., Hughes, K.H., Martelli, F., and Burghardt, I. (2010). Effective spectral densities for system-environment dynamics at conical intersections: $S_2 - S_1$ conical intersection in pyrazine. *Chem. Phys.* 377: 21.
86 Bonfanti, M., Hughes, K.H., Burghardt, I., and Martinazzo, R. (2015). Vibrational relaxation and decoherence in structured environments: a numerical investigation. *Ann. Phys.* 527: 556.
87 Hughes, K.H., Christ, C.D., and Burghardt, I. (2009). Effective-mode representation of non-Markovian dynamics: a hierarchical approximation of the spectral density. II. Application to environment-induced non-adiabatic dynamics. *J. Chem. Phys.* 131: 124108.
88 Parandekar, P.V. and Tully, J.C. (2005). Mixed quantum-classical equilibrium. *J. Chem. Phys.* 122: 094102.

89 Parandekar, P.V. and Tully, J.C. (2006). Detailed balance in Ehrenfest mixed quantum-classical dynamics. *J. Chem. Theory Comput.* 2: 229.

90 Koch, W. and Zhang, D.H. (2014). Separable potential energy surfaces from multiplicative artificial neural networks. *J. Chem. Phys.* 141: 21101.

91 Koch, W., Bonfanti, M., Eisenbrandt, P. et al. (2019). Two-layer Gaussian-based MCTDH study of the $S_2 - S_1$ vibronic absorption spectrum of formaldehyde using multiplicative neural network potentials. *J. Chem. Phys.* 151: 064121.

92 Richings, G.W. and Habershon, S. (2018). MCTDH on-the-fly: efficient grid-based quantum dynamics without pre-computed potential energy surfaces. *J. Chem. Phys.* 148: 134116.

93 Worth, G.A. (2020). Quantics: a general purpose package for quantum molecular dynamics simulations. *Comput. Phys. Commun.* 248: 107040.

94 Martinazzo, R. and Burghardt, I. (2020). Local-in-time error in variational quantum dynamics. *Phys. Rev. Lett.* 124: 150601.

13

Gaussian Wave Packets and the DD-vMCG Approach

Graham A. Worth[1] *and Benjamin Lasorne*[2]

[1] *University College London, United Kingdom*
[2] *Université de Montpellier, France*

Abstract

This chapter provides an overview of the *variational multi-configurational Gaussian* (vMCG) approach, which belongs to the family of *Gaussian wave packet* (GWP) methods. It is based on similar grounds to the *spawning* and related approaches (see Chapter 14), but has explicit coupling between the GWP quantum amplitudes. This means that it relies on "quantum trajectories" followed by the centers of the GWPs, rather than classical ones. By this, it is to be understood that the mean positions and momenta are specifically obtained from equations of motions derived from a variational solution to the *time-dependent Schrödinger equation* (TDSE), in addition to the expansion coefficients, thus resulting in a favorable convergence on the exact solution. GWP methods are attractive ways to solve the TDSE, alternative to grid-based methods such as the *multi-configurational time-dependent Hartree* (MCTDH) approach (see Chapter 12). Their main advantage is that they provide a way to run direct dynamics with quantum nuclei, whereby the potential-energy surfaces and non-adiabatic couplings are calculated on-the-fly, much as in trajectory-based approaches (see Chapters 13–17). This is achieved via an interface between quantum-dynamics and quantum-chemistry programs. The *direct dynamics* variant of vMCG is known as DD-vMCG. The implementation of the method and its performance will be illustrated with three examples: salicylaldimine to show phase-space coverage and tunneling, the butatriene cation to demonstrate how it works within a non-adiabatic context, and formamide to exemplify direct dynamics.

13.1 Historical Background

Using *Gaussian wave packets* (GWPs) for solving the *time-dependent Schrödinger equation* (TDSE) and simulating quantum nuclear dynamics within molecular systems dates back to the seminal work of Heller [1, 2]. Such basis functions provide a link between the delocalized description of quantum mechanics with wave packets and the localized, trajectory-based, description of classical mechanics. GWP-based methods thus provide a way to run direct quantum dynamics, calculating the potential-energy surfaces and non-adiabatic couplings on-the-fly, thus bypassing a big bottleneck to practical simulations while still including all quantum effects.

Quantum Chemistry and Dynamics of Excited States: Methods and Applications,
First Edition. Edited by Leticia González and Roland Lindh.

Originally, Heller used a single GWP, which in one dimension (coordinate x) has the form [1]

$$g(x,t) = \exp\left[\frac{\mathrm{i}}{\hbar}(\alpha(t)(x-q(t))^2 + p(t)(x-q(t)) + \gamma(t))\right] , \tag{13.1}$$

to approximate the time evolving wave packet. In this, $\alpha(t)$, proportional to the squared reciprocal width, the mean position and momentum, $q(t)$ and $p(t)$, and the complex phase, $\gamma(t)$, are all time-dependent parameters. It was shown that in a solution of the TDSE, $\{q(t), p(t)\}$ follow a classical trajectory in phase space, $\alpha(t)$ oscillates with time, while the real part of $\gamma(t)$ carries the quantum information along the trajectory and is known as the action (its imaginary part is related to the norm of the wave packet).

Extending this to a multi-dimensional case, when the reciprocal width becomes a matrix with off-diagonal elements carrying the correlation between nuclear coordinates, is straightforward [3]. It is a standard result of quantum mechanics to show that in a harmonic potential energy a time-evolving GWP is the exact solution, and it is assumed to be a reasonable solution in other cases when the potential energy is changing slowly compared to the width. This approximation has been used in a number of studies, but gives good results only for simple systems and very short timescales. The single *thawed* GWP ansatz, with time-dependent width, is simply not flexible enough to follow a real molecular system.

To improve this, Heller introduced the idea of describing the wave packet by a superposition of *frozen* GWPs with constant widths [2],

$$\Psi(\mathbf{x},t) = \sum_j g_j(\mathbf{x},t) , \tag{13.2}$$

each following a different classical trajectory. Much work has been done to make the frozen Gaussian superposition a useful, general method by connection to semiclassical approaches [4–6] and the use of a variational *minimum error method* [7–9]. Despite the promise of a simple scheme, however, GWP propagation was found to be numerically very unstable.

A resurgence of interest, continuing to the present day, was driven by the development of the *spawning* algorithm of Martínez, Ben-Nun, and Levine [10, 11] (see Chapter 14). This uses an expansion of frozen GWPs with redundant coefficients,

$$\Psi(\mathbf{x},t) = \sum_j A_j(t) g_j(\mathbf{x},t) . \tag{13.3}$$

The GWPs still follow classical trajectories, but the coefficients evolve so as to variationally solve the TDSE within the GWP basis set, much as in any multi-configurational time-dependent method. The redundant coefficients are helpful for the stability of the method, as the GWPs can be kept as normalized functions with constrained phases.

Spawning has pioneered direct quantum dynamics whereby potential-energy surfaces and non-adiabatic couplings are provided on-the-fly [12, 13] via an interface to quantum chemistry. In this form, known as *ab initio multiple spawning* (AIMS), the spawning method and other related ones (see Chapter 14) have become powerful tools for simulating non-adiabatic dynamics in photoexcited molecules.

The subject of the present chapter is a method termed *variational multi-configurational Gaussian* (vMCG), with its *direct dynamics* variant being known as DD-vMCG, in which the GWPs no longer follow classical trajectories. Instead, it stems from the *multi-configurational time-dependent Hartree* (MCTDH) method (see Chapter 12), where a fully variational solution to the TDSE is derived using a time-dependent basis set of low-dimensional functions for different parts of the system which are coupled through mean-field operators. The *Gaussian* MCTDH (G-MCTDH) scheme uses GWPs as

the basis functions, either for all or a subset of degrees of freedom. If multi-dimensional GWP basis functions are used that include all degrees of freedom of the system, the approach boils down to a GWP-based method, namely vMCG. As a result of the fully variational development, the vMCG GWPs do not follow classical trajectories, and it is thus expected that fewer basis functions are required for a given convergence threshold.

A description of the vMCG and DD-vMCG methods is given in a number of the original papers as their development proceeded [14–24]. Here we summarize the general formalism of the vMCG method and illustrate the principles of its current *direct dynamics* implementation in the Quantics quantum-dynamics package [25] with three examples: salicylaldimine to show phase-space coverage and tunneling, the butatriene cation to demonstrate how it works within a non-adiabatic context, and formamide to exemplify direct dynamics.

13.2 Basic Theory

13.2.1 Gaussian Wave Packets

Before entering into the details of the vMCG machinery, it is useful to recall some basic properties of GWPs expressed in various ways. A one-dimensional description is sufficient at the moment in order to grasp the essentials.

Let us first consider a normalized Gaussian distribution, function of x, with centre q (mean position) and variance σ (standard deviation or width),

$$|g(x)|^2 = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-q}{\sigma}\right)^2\right) , \tag{13.4}$$

where

$$\int_{-\infty}^{+\infty} |g(x)|^2 \mathrm{d}x = 1 . \tag{13.5}$$

In quantum-mechanical terms, $|g(x)|^2$ stands for a density of probability within the x-space. The corresponding amplitude of probability is the following wave function,

$$g(x) = \frac{1}{\sqrt{\sigma\sqrt{2\pi}}} \exp\left(-\frac{1}{4}\left(\frac{x-q}{\sigma}\right)^2 + \mathrm{i}\phi(x)\right) , \tag{13.6}$$

where $\phi(x)$ is a real-valued phase function, yet to be specified.

The ground eigenstate of the following harmonic Hamiltonian operator,

$$\hat{H} = -\frac{\hbar^2}{2\mu}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{k}{2}(x-c)^2 , \tag{13.7}$$

reads (conventionally taken with a unity phase factor)

$$g_0(x) = \left(\frac{\mu\omega}{\pi\hbar}\right)^{1/4} \exp\left(-\frac{\mu\omega(x-c)^2}{2\hbar}\right) , \tag{13.8}$$

where $\omega = \sqrt{\frac{k}{\mu}}$. The natural unit of length[1] of this system is $\beta = \sqrt{\frac{\hbar}{\mu\omega}}$.

1 The shifted and frequency-mass-weighted coordinate $X = (x-c)/\beta$ is dimensionless, and the corresponding ground eigenstate reads $G_0(X) = \pi^{-1/4} \exp\left(-\frac{X^2}{2}\right)$. Its standard deviation is $\frac{1}{\sqrt{2}}$.

If the arbitrary phase factor $e^{i\phi(x)}$ is a constant, the two Gaussian functions, $g(x)$ and $g_0(x)$ match (up to the arbitrary phase), provided both $q = c$ and $\sigma = \frac{\beta}{\sqrt{2}}$. Let us now consider the time evolution of a GWP, $g(x,t)$, such that $g(x,t=0) \equiv g(x)$. If $q < c$, it will move to the right, and to the left if $q > c$. The mean position will then oscillate at the fundamental angular frequency, ω. Now, if $\sigma < \frac{\beta}{\sqrt{2}}$, it will start spreading, or contracting if $\sigma > \frac{\beta}{\sqrt{2}}$, and the standard deviation will then oscillate at twice the fundamental angular frequency, 2ω. Meanwhile, the wave packet will acquire a finite mean momentum, hence a group velocity, which also will oscillate at ω. In addition, the density of probability will stay a Gaussian function over time. Such a quasiclassical behavior for a GWP within a harmonic well is a well-known textbook result of quantum mechanics (see also the work of Heller [1, 2]).

The general expression of the normalized GWP is thus

$$g(x,t) = \frac{1}{\sqrt{\sigma(t)\sqrt{2\pi}}} \exp\left(-\frac{1}{4}\left(\frac{x-q(t)}{\sigma(t)}\right)^2 + i\frac{p(t)(x-q(t))}{\hbar} + i\theta(t) \right) , \tag{13.9}$$

where $p(t)$ is the mean momentum and $\theta(t)$ a real phase offset. In a semi-classical context $\hbar\theta(t)$ is known as the action. Note that the phase function now reads $\phi(x,t) = \frac{p(t)(x-q(t))}{\hbar} + \theta(t)$.

Now, let us make the connection with the usual formulation used in the literature on vMCG. Considering the exponent as a quadratic form yields

$$g(x,t) = \exp(\zeta(t)x^2 + \xi(t)x + \eta(t)) . \tag{13.10}$$

Note that the normalization factor has been exponentiated for convenience.[2] Identification with Eq. (13.9) yields

$$\zeta(t) = -\frac{1}{4\sigma(t)^2} , \tag{13.11}$$

$$\xi(t) = \frac{q(t)}{2\sigma(t)^2} + i\frac{p(t)}{\hbar} , \tag{13.12}$$

$$\eta(t) = -\frac{q(t)^2}{4\sigma(t)^2} - i\frac{p(t)q(t)}{\hbar} + i\theta(t) - \frac{\ln(\sigma(t)\sqrt{2\pi})}{2} . \tag{13.13}$$

Using Heller's form, as in Eq. (13.1), we would have

$$\eta(t) = \frac{i}{\hbar}(\alpha(t)q(t)^2 - p(t)q(t) + \gamma(t)) , \tag{13.14}$$

with the complex phase

$$\frac{\gamma(t)}{\hbar} = \theta(t) + i\frac{\ln(\sigma(t)\sqrt{2\pi})}{2} . \tag{13.15}$$

The real part of this term corresponds to a phase convention, and the imaginary part to a normalization constraint.

From now on, we shall restrict our discussion to *frozen* GWPs for which σ is a constant with respect to time. The normalization factor is thus a constant too, and its contribution to the real part of $\eta(t)$ can be safely removed from its definition and rather used explicitly as a prefactor (this

2 Exponentiation of the normalization factor is formally and numerically acceptable, but not physically because it is a dimensioned quantity, the unit of which will not appear to a unique order in the power series. More rigorously, we should first divide σ by the working unit of length, [L], thus getting $\frac{\ln(\sigma\sqrt{2\pi}/[L])}{2}$ into the exponent, and multiply the exponential by the prefactor $[L^{-1/2}]$.

will be of no consequence on the expression of the time derivative of $\eta(t)$). Now, the real-valued function $\theta(t)$ appears as a gauge angle that remains to be determined upon using some conventional constraint. This redundancy is typical of a multi-configurational wave function ansatz, where some or all of the phase can be incorporated into the expansion coefficients in Eq. (13.3). In the current implementation of vMCG we have chosen $\theta(t) \equiv 0$, which seems to often result in a more stable propagation.

In what follows, we will thus use the two following GWP forms,

$$g(x,t) = \frac{1}{\sqrt{\sigma\sqrt{2\pi}}} \exp\left(-\frac{(x-q(t))^2}{4\sigma^2} + \mathrm{i}\frac{p(t)(x-q(t))}{\hbar}\right) , \tag{13.16}$$

or

$$g(x,t) = \frac{1}{\sqrt{\sigma\sqrt{2\pi}}} \exp\left(-\frac{x^2}{4\sigma^2} + \xi(t)x + \eta(t)\right) , \tag{13.17}$$

with

$$\xi(t) = \frac{q(t)}{2\sigma^2} + \mathrm{i}\frac{p(t)}{\hbar} , \tag{13.18}$$

$$\eta(t) = -\frac{q(t)^2}{4\sigma^2} - \mathrm{i}\frac{p(t)q(t)}{\hbar} . \tag{13.19}$$

Such a GWP thus appears as a normalized Gaussian envelope centered on $q(t)$ multiplied by a Fourier basis function also centered on $q(t)$. Other choices for $\theta(t)$ may result in the Fourier function centered elsewhere. Only the complex, first-order parameter, $\xi(t)$, must be varied with time explicitly. This is equivalent to getting the time evolution of the real phase-space parameters, $q(t)$ and $p(t)$. The complex, zeroth-order parameter, $\eta(t)$ can be reconstructed at any time from the knowledge of the former two.

Now, let us consider a basis set of GWPs, $g_j(x,t)$, each determined by its own $\{p_j(t), q_j(t)\}$ or $\xi_j(t)$. As will be made clearer later on, Gaussian moments are fundamental quantities for expressing Gaussian matrix elements. They can be obtained from the moment generator as follows (omitting the dependence on time for notational simplicity),

$$M_{jk}(u) = \langle g_j | \mathrm{e}^{u\hat{x}} | g_k \rangle = \exp\left(\frac{\sigma^2}{2}u^2 + \left(\frac{q_j+q_k}{2}\right)u - \mathrm{i}\frac{(p_j-p_k)\sigma^2}{\hbar}u\right) S_{jk} , \tag{13.20}$$

where the overlap can be expressed as

$$S_{jk} = \langle g_j | g_k \rangle = \exp\left(-\frac{\left(\frac{q_j-q_k}{2}\right)^2}{2\sigma^2} - \frac{(p_j-p_k)^2\sigma^2}{2\hbar^2} + \mathrm{i}\frac{(p_j+p_k)\left(\frac{q_j-q_k}{2}\right)}{\hbar}\right) . \tag{13.21}$$

Differentiation once and twice with respect to u yields

$$M'_{jk}(u) = \langle g_j | \hat{x}\mathrm{e}^{u\hat{x}} | g_k \rangle = \left(\sigma^2 u + \left(\frac{q_j+q_k}{2}\right) - \mathrm{i}\frac{(p_j-p_k)\sigma^2}{\hbar}\right) M_{jk}(u) , \tag{13.22}$$

$$M''_{jk}(u) = \langle g_j | \hat{x}^2\mathrm{e}^{u\hat{x}} | g_k \rangle = \left(\sigma^2 + \left(\sigma^2 u + \left(\frac{q_j+q_k}{2}\right) - \mathrm{i}\frac{(p_j-p_k)\sigma^2}{\hbar}\right)^2\right) M_{jk}(u) . \tag{13.23}$$

The first three Gaussian moments are obtained upon setting $u = 0$,

$$M_{jk}^{(0)} = M_{jk}(0) = \langle g_j|\hat{1}|g_k\rangle = S_{jk} \ , \tag{13.24}$$

$$M_{jk}^{(1)} = M'_{jk}(0) = \langle g_j|\hat{x}|g_k\rangle = \sigma^2(\xi_j^* + \xi_k)S_{jk} \ , \tag{13.25}$$

$$M_{jk}^{(2)} = M''_{jk}(0) = \langle g_j|\hat{x}^2|g_k\rangle = \sigma^2(1 + \sigma^2(\xi_j^* + \xi_k)^2)S_{jk} \ . \tag{13.26}$$

13.2.2 General Equations of Motion

13.2.2.1 Coefficients and Parameters

The vMCG *equations of motion* (EOMs) are obtained, consistent with the MCTDH approach, by applying the Dirac–Frenkel variational principle [26, 27] to the TDSE,

$$\langle \delta\Psi(t)|\hat{H} - \mathrm{i}\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle = 0 \ , \tag{13.27}$$

guaranteeing the optimal evolution (minimal error) of the approximate nuclear wave packet.

The following derivation is general and applies to any sum of time-dependent, parametric basis functions, multiplied by time-dependent expansion coefficients,

$$|\Psi(t)\rangle = \sum_j A_j(t)|g_j(t)\rangle \ . \tag{13.28}$$

Each $|g_j(t)\rangle$ is a given functional form, fully determined by a set of complex, time-dependent parameters $\{\cdots, \lambda_j(t), \cdots, \nu_j(t), \cdots\}$. The parametric basis set is *a priori* non-orthogonal. At this stage, it does not have to be Gaussian.

Applying the Dirac–Frenkel variational principle (Eq. (13.27)) to the wave packet ansatz (Eq. (13.28)), two coupled EOMs are obtained: one for the time-dependent coefficients, $\mathbf{A}(t) = \{\cdots, A_j(t), \cdots, A_k(t), \cdots\}$, and one for the time-dependent parameters, $\mathbf{\Lambda}(t) = \{\cdots, \lambda_j(t), \cdots, \nu_j(t), \cdots, \lambda_k(t), \cdots, \nu_k(t), \cdots\}$.

Varying both coefficients and parameters yields

$$\mathrm{i}\hbar\sum_k S_{jk}\frac{\partial A_k}{\partial t} + \mathrm{i}\hbar\sum_k\sum_\nu S_{jk}^{(0\nu)}\frac{\partial \nu_k}{\partial t}A_k = \sum_k H_{jk}A_k \ , \tag{13.29}$$

$$\mathrm{i}\hbar\sum_k S_{jk}^{(\lambda 0)}A_j^*\frac{\partial A_k}{\partial t} + \mathrm{i}\hbar\sum_k\sum_\nu S_{jk}^{(\lambda\nu)}\frac{\partial \nu_k}{\partial t}A_j^*A_k = \sum_k H_{jk}^{(\lambda 0)}A_j^*A_k \ , \tag{13.30}$$

where S_{jk} is the overlap matrix element,

$$S_{jk} = \langle g_j|g_k\rangle \ , \tag{13.31}$$

H_{jk} is the Hamiltonian matrix element,

$$H_{jk} = \langle g_j|\hat{H}|g_k\rangle \ , \tag{13.32}$$

and additional definitions are

$$S_{jk}^{(0\nu)} = \left\langle g_j\middle|\frac{\partial g_k}{\partial \nu_k}\right\rangle \ ; \ S_{jk}^{(\lambda 0)} = \left\langle \frac{\partial g_j}{\partial \lambda_j}\middle|g_k\right\rangle \ ; \ S_{jk}^{(\lambda\nu)} = \left\langle \frac{\partial g_j}{\partial \lambda_j}\middle|\frac{\partial g_k}{\partial \nu_k}\right\rangle \ , \tag{13.33a}$$

$$H_{jk}^{(\lambda 0)} = \left\langle \frac{\partial g_j}{\partial \lambda_j}|\hat{H}|g_k\right\rangle \ . \tag{13.33b}$$

Note that $\left\langle \frac{\partial g_j}{\partial \lambda_j} \right|$ means that g_j^* is complex-differentiated with respect to the complex-conjugate parameter λ_j^* (whereas g_k is complex-differentiated with respect to the original parameter ν_k in $\left| \frac{\partial g_k}{\partial \nu_k} \right\rangle$).

Now, we introduce the density matrix element for the expansion coefficients,

$$\rho_{jk} = A_j^* A_k \ , \tag{13.34}$$

and the overlap time-derivative matrix element,

$$\tau_{jk} = \left\langle g_j \middle| \frac{\partial g_k}{\partial t} \right\rangle = \sum_\nu S_{jk}^{(0\nu)} \frac{\partial \nu_k}{\partial t} \ . \tag{13.35}$$

The diagonal of this matrix must be purely imaginary for the basis functions to stay normalized during the propagation. In addition, various phase conventions can be applied.

The EOM for the time-dependent coefficients, Eq. (13.29), can be recast as

$$\mathrm{i}\hbar \frac{\partial \mathbf{A}}{\partial t} = \mathbf{S}^{-1}(\mathbf{H} - \mathrm{i}\hbar\boldsymbol{\tau})\mathbf{A} \ , \tag{13.36}$$

which is typical of a standard coefficient-EOM, except for the basis set not being orthogonal which results in the presence of the inverse of $\mathbf{S}$. Note that $\mathrm{i}\hbar\boldsymbol{\tau}$ plays a similar role to the matrix of the phase-constraint operator in MCTDH EOMs, see Chapter 12, and requires a real diagonal for the basis functions to stay normalized. This equation is the same as the one for the coefficient time evolution in G-MCTDH (See Eq. (12.28) in Chapter 12) and in spawning (See Chapter 14).

Injecting Eq. (13.36) into Eq. (13.30) and regrouping terms involving $\frac{\partial \nu_k}{\partial t}$ yields

$$\mathrm{i}\hbar \sum_k \sum_\nu \rho_{jk} (S_{jk}^{(\lambda\nu)} - [\mathbf{S}^{(\lambda 0)}\mathbf{S}^{-1}\mathbf{S}^{(0\nu)}]_{jk}) \frac{\partial \nu_k}{\partial t} = \sum_k \rho_{jk} (H_{jk}^{(\lambda 0)} - [\mathbf{S}^{(\lambda 0)}\mathbf{S}^{-1}\mathbf{H}]_{jk}) \ . \tag{13.37}$$

The parameter-EOM thus gets a compact matrix form [14],

$$\mathrm{i}\hbar \frac{\partial \mathbf{\Lambda}}{\partial t} = \mathbf{C}^{-1}\mathbf{Y} \ , \tag{13.38}$$

where the matrix $\mathbf{C}$, and vectors $\mathbf{Y}$ and $\mathbf{\Lambda}$ are defined as

$$C_{j\lambda,k\nu} = \rho_{jk} (S_{jk}^{(\lambda\nu)} - [\mathbf{S}^{(\lambda 0)}\mathbf{S}^{-1}\mathbf{S}^{(0\nu)}]_{jk}) \ , \tag{13.39a}$$

$$Y_{j\lambda} = \sum_k \rho_{jk} (H_{jk}^{(\lambda 0)} - [\mathbf{S}^{(\lambda 0)}\mathbf{S}^{-1}\mathbf{H}]_{jk}) \ , \tag{13.39b}$$

$$\Lambda_{j\lambda} = \lambda_j \ . \tag{13.39c}$$

Note that both $\mathbf{C}$ and $\mathbf{Y}$ are functionals of $\mathbf{\Lambda}$ via the various overlap/Hamiltonian-type matrices involved. They also are functionals of $\mathbf{A}$ via the density matrix. Similarly, the right-hand side of the coefficient-EOM, Eq. (13.36), is a functional of both $\mathbf{\Lambda}$ and $\mathbf{A}$. We thus have coupled, first-order differential equations with respect to time that must be solved simultaneously. Details on integration schemes can be found elsewhere, for example in Ref. [23] and references therein.

13.2.2.2 CX-Formalism

Solving the parameter-EOM involves inversion of the matrix $\mathbf{C}$, which may cause numerical difficulties. A reformulation of the EOM aims at minimizing potential sources of error due to this.

Examining the respective expressions of $\mathbf{C}$ (Eq. (13.39a)) and $\mathbf{Y}$ (Eq. (13.39b)) reveals a similar structure, which can be exploited if Hamiltonian-type matrices can be expressed in terms of overlap-type matrices. First, let us separate $\hat{H}$ into two parts,

$$\hat{H} = \hat{H}^0 + \hat{H}^R \,, \tag{13.40}$$

where the zeroth-order Hamiltonian operator is specifically defined such that

$$\hat{H}^0|g_k\rangle = X_k^{(0)}|g_k\rangle + \sum_\nu X_k^{(\nu)} \left|\frac{\partial g_k}{\partial \nu_k}\right\rangle \,, \tag{13.41}$$

with X-coefficients that are to be determined according to the specific problem at hand, and $\hat{H}^R$ is the remaining part of the total Hamiltonian operator. We then get

$$H_{jk} = S_{jk}X_k^{(0)} + \sum_\nu X_k^{(\nu)} S_{jk}^{(0\nu)} + H_{jk}^R \,, \tag{13.42a}$$

$$H_{jk}^{(\lambda 0)} = S_{jk}^{(\lambda 0)}X_k^{(0)} + \sum_\nu X_k^{(\nu)} S_{jk}^{(\lambda\nu)} + H_{jk}^{R(\lambda 0)} \,, \tag{13.42b}$$

such that $\mathbf{Y}$ can in turn be separated into two parts,

$$\mathbf{Y} = \mathbf{Y}^0 + \mathbf{Y}^R \,, \tag{13.43}$$

where[3]

$$Y_{j\lambda}^0 = \sum_k \sum_\nu C_{j\lambda,k\nu} X_k^{(\nu)} \,. \tag{13.44}$$

This leads to a partial decoupling of the parameter-EOM [14, 18],

$$\mathrm{i}\hbar \frac{\partial \mathbf{\Lambda}}{\partial t} = \mathbf{X} + \mathbf{C}^{-1}\mathbf{Y}^R \,. \tag{13.45}$$

The contribution due to $\mathbf{X}$ induces uncoupled time evolution of the GWP parameters.

The so-called CX-formalism has advantages. First, it improves the stability of the propagation upon decreasing the possible numerical error due to the $\mathbf{C}$-matrix inversion, especially if the "residual" contribution is small compared to that due to $\mathbf{X}$. Second, as described in subsection 13.2.3, it allows a division of the parameter-EOM into "classical" and "non-classical" parts. A more rigorous discussion of the classical limit of vMCG is given in Chapter 12 and Ref. [28].

13.2.2.3 Nuclear and Electronic Degrees of Freedom

Let us now separate explicitly the nuclear (coordinates $\mathbf{x}$) and electronic (index s, with a parametric dependence on $\mathbf{x}$) degrees of freedom. The electronic basis set is considered orthonormal,

$$\forall \mathbf{x}, \quad \langle s;\mathbf{x}|r;\mathbf{x}\rangle = \delta_{sr} \,. \tag{13.46}$$

In vMCG, as in MCTDH, there are two alternative formulations for non-adiabatic problems that involve multiple electronic states. In the *multi-set* formalism, a different set of GWPs is used for each electronic state [18],

$$|\Psi(\mathbf{x},t)\rangle = \sum_s \sum_j A_j^{(s)}(t) g_j^{(s)}(\mathbf{x},t)|s;\mathbf{x}\rangle \,. \tag{13.47}$$

---

3 Contributions involving $X_k^{(0)}$ (see Eqs. (13.42a) and (13.42b)) do not affect $\mathbf{Y}^0$, hence do not enter the parameter-EOM, since the corresponding term in $H_{jk}^{(\lambda 0)} - [\mathbf{S}^{(\lambda 0)}\mathbf{S}^{-1}\mathbf{H}]_{jk}$ is $S_{jk}^{(\lambda 0)}X_k^{(0)} - [\mathbf{S}^{(\lambda 0)}\mathbf{S}^{-1}\mathbf{S}]_{jk}X_k^{(0)} = 0$. However, they do not disappear from the coefficient-EOM and may induce uncoupled phase shifts within the expansion coefficients: $\mathrm{i}\hbar\frac{\partial A_j}{\partial t} = X_j^{(0)}A_j + \sum_k (\cdots)_{jk}A_k$.

In contrast, in the *single-set* formalism, a unique and common set of nuclear basis functions is used, which implies that the electronic states are included as an extra degree of freedom described by its own finite basis set[4],

$$|\Psi(\mathbf{x},t)\rangle = \sum_j \left(\sum_s A_j^{(s)}(t)|s;\mathbf{x}\rangle \right) g_j(\mathbf{x},t) \ . \tag{13.48}$$

This is similar to an Ehrenfest approach: each nuclear wave packet is multiplied by a time-dependent, weighted sum of electronic states, $\sum_s A_j^{(s)}(t)|s;\mathbf{x}\rangle$, which generates a time-dependent, average potential energy.

In the *multi-set* formalism, the EOMs are obtained from the same way as the general ones, simply upon replacing $|g_j(t)\rangle$ by $g_j(\mathbf{x},t)|s;\mathbf{x}\rangle$ and using the orthonormality of the electronic basis set. Overlap-type matrices thus have a single electronic index, while Hamiltonian-type matrices have two electronic indices, able to couple nuclear basis functions on different electronic states. For the coefficients, we get

$$\mathrm{i}\hbar \frac{\partial \mathbf{A}^{(s)}}{\partial t} = \mathbf{S}^{(s)-1}(\mathbf{H}^{(ss)} - \mathrm{i}\hbar\boldsymbol{\tau}^{(s)})\mathbf{A}^{(s)} + \sum_{r\neq s} \mathbf{H}^{(sr)}\mathbf{A}^{(r)} \ , \tag{13.49}$$

and for the parameters,

$$\mathrm{i}\hbar \frac{\partial \boldsymbol{\Lambda}^{(s)}}{\partial t} = \mathbf{C}^{(s)-1}\mathbf{Y}^{(s)} \ , \tag{13.50}$$

with

$$C^{(s)}_{j\lambda,k\nu} = \rho^{(ss)}_{jk}(S^{(s,\lambda\nu)}_{jk} - [\mathbf{S}^{(s,\lambda 0)}\mathbf{S}^{(s)-1}\mathbf{S}^{(s,0\nu)}]_{jk}) \ , \tag{13.51a}$$

$$Y^{(s)}_{j\lambda} = \sum_r \sum_k \rho^{(sr)}_{jk}(H^{(sr,\lambda 0)}_{jk} - [\mathbf{S}^{(s,\lambda 0)}\mathbf{S}^{(s)-1}\mathbf{H}^{(sr)}]_{jk}) \ . \tag{13.51b}$$

In the *single-set* formalism, it must be understood that a variation of a parameter λ_j produces a term involving a weighted sum over all electronic states, $\left\langle \frac{\partial g_j}{\partial \lambda_j} \right| \left(\sum_s A_j^{(s)*}(t)\langle s;\mathbf{x}| \right)$ in lieu of the single, (s,j)-composite-indexed term, $\left\langle \frac{\partial g_j^{(s)}}{\partial \lambda_j} \right| \langle s;\mathbf{x}|A_j^{(s)*}(t)$ that occurs in the *multi-set* formalism. In other words, the *single-set* formalism involves an extra contraction over the electronic index. Overlap-type matrices no longer depend on the electronic index, but the Hamiltonian-type matrices keep two electronic indices, and are still able to couple nuclear basis functions on different electronic states. For the coefficients, this yields

$$\mathrm{i}\hbar \frac{\partial \mathbf{A}^{(s)}}{\partial t} = \mathbf{S}^{-1}(\mathbf{H}^{(ss)} - \mathrm{i}\hbar\boldsymbol{\tau})\mathbf{A}^{(s)} + \sum_{r\neq s} \mathbf{H}^{(sr)}\mathbf{A}^{(r)} \ , \tag{13.52}$$

and for the parameters,

$$\mathrm{i}\hbar \frac{\partial \boldsymbol{\Lambda}}{\partial t} = \mathbf{C}^{-1}\mathbf{Y} \ , \tag{13.53}$$

4 Within the MCTDH framework, the *multi-set* formalism corresponds to single particles combining nuclear and electronic degrees of freedom together, while the *single-set* formalism involves direct products of two types of single particles – nuclear and electronic – with repeated indices: $\{\cdots, g_j|s\rangle, \cdots, g_k|s\rangle, \cdots, g_j|r\rangle, \cdots, g_k|r\rangle, \cdots\}$. Contracting over electronic degrees of freedom thus creates mean fields for the nuclear basis functions.

with

$$C_{j\lambda,k\nu} = \sum_s \rho_{jk}^{(ss)} (S_{jk}^{(\lambda\nu)} - [\mathbf{S}^{(\lambda 0)}\mathbf{S}^{-1}\mathbf{S}^{(0\nu)}]_{jk}) \,, \tag{13.54a}$$

$$Y_{j\lambda} = \sum_s \sum_r \sum_k \rho_{jk}^{(sr)} (H_{jk}^{(sr,\lambda 0)} - [\mathbf{S}^{(\lambda 0)}\mathbf{S}^{-1}\mathbf{H}^{(sr)}]_{jk}) \,. \tag{13.54b}$$

In general it is found that the *multi-set* formalism is preferred in MCTDH calculations, as the basis functions are able to adapt better to the different states so the number of expansion coefficients is minimized. However, it does require more basis functions, which for direct dynamics may be a crucial expense, as more quantum-chemistry calculations may be required. In addition, basis functions with very small coefficients may follow erratic trajectories, which may cause numerical issues, especially for direct dynamics. DD-vMCG calculations are thus rather run within the *single-set* formalism.

13.2.3 Variational Multi-Configurational Gaussian Approach

Here, we provide some additional specifics about the EOMs when a GWP functional form is used for the basis functions. The vMCG ansatz is given by Eqs. (13.3), (13.47), or (13.48), assuming that $g_j(\mathbf{x}, t)$ or $g_j^{(s)}(\mathbf{x}, t)$ are GWPs; see Eqs. (13.16) or (13.17) in the one-dimensional case exposed in subsection 13.2.1. As already mentioned, we shall not focus in this chapter on thawed GWPs, as they have proved to yield numerical instabilities. Frozen GWPs are given in Eqs. (13.9) or (13.10), with an explicit normalization constraint and the most usual phase convention.

For several degrees of freedom and in the separable case (diagonal width matrix), each multi-dimensional GWP is a product of one-dimensional GWPs. However, it must be understood that there is no redundancy of one-dimensional basis functions among the set of multi-dimensional basis functions. In other words, each basis function corresponds to a *single particle* made of all degrees of freedom combined together.[5]

For one-dimensional, frozen GWPs, the complex, time-dependent parameters are $\{\eta_j, \xi_j\}$. They can be reconstructed from $\{q_j, p_j\}$, from Eqs. (13.18) and (13.19) (note that various phase conventions can be used for η_j). In the multi-dimensional case, η_j is still a scalar, but $\boldsymbol{\xi}_j$ is a vector the size of $\mathbf{x}$, and the linear term in the exponent of the GWP is the scalar product of $\boldsymbol{\xi}_j$ with $\mathbf{x}$.

Let us now consider a typical harmonic Hamiltonian (see Eq. (13.7)), expanded around $x = q_k$,

$$\hat{H} = -\frac{\hbar^2}{2\mu_k}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V_k + V_k'(x - q_k) + \frac{1}{2}V_k''(x - q_k)^2 \,, \tag{13.55}$$

where V_k, V_k', and V_k'' are the potential energy and its gradient and Hessian at q_k. If q_k is specifically taken as the centre of the GWP $g_k(x, t)$, this is called a *local harmonic approximation* (LHA). Each GWP "feels" its own local quadratic potential. As q_k depends on time, we can consider that the centre of the GWP follows a "quantum trajectory", along which the potential energy and its first and second derivatives can be calculated on-the-fly, as in any classical or semiclassical trajectory-based direct dynamics. This is the bottom line of the DD-vMCG approach and will be further illustrated in the next section.

---

5 Let us consider four two-dimensional GWPs: $\{g_1(x, y, t), g_2(x, y, t), g_3(x, y, t), g_4(x, y, t)\}$. Assuming one-dimensional single particles would imply that the basis set is a direct product of, for example, two one-dimensional GWP per mode: $\{g_{1_x}^{(x)}(x, t), g_{2_x}^{(x)}(x, t)\} \bigotimes \{g_{1_y}^{(y)}(y, t), g_{2_y}^{(y)}(y, t)\}$, whereby $\{1, 2, 3, 4\} \equiv \{1_x1_y, 1_x2_y, 2_x1_y, 2_x2_y\}$. Doing so, the centers of the four GWPs would then be constrained to be at the vertices of a rectangle within the (x, y)-plane.

Each LHA Hamiltonian matrix element can be expanded in terms of the first three Gaussian moments (Eqs. (13.24), (13.25), and (13.26)),

$$H_{jk} = M_{jk}^{(0)} X_k^{(0)} + M_{jk}^{(1)} X_k^{(1)} + M_{jk}^{(2)} X_k^{(2)} \,, \tag{13.56}$$

where the X-coefficients in Eq. (13.56) are as follows,

$$X_k^{(0)} = -\frac{\hbar^2}{\mu_k}\left(-\frac{1}{4\sigma^2} + \frac{\xi_k^2}{2}\right) + V_k - V_k' q_k + \frac{1}{2} V_k'' q_k^2 \,, \tag{13.57a}$$

$$X_k^{(1)} = \frac{\hbar^2}{\mu_k}\frac{\xi_k}{2\sigma^2} + V_k' - V_k'' q_k \,, \tag{13.57b}$$

$$X_k^{(2)} = -\frac{\hbar^2}{\mu_k}\frac{1}{8\sigma^4} + \frac{1}{2} V_k'' \,. \tag{13.57c}$$

Note that the first-order coefficients can be recast as

$$X_k^{(1)} = \mathrm{i}\hbar \frac{1}{2\sigma^2}\frac{p_k}{\mu_k} + V_k' - 2X_k^{(2)} q_k \,. \tag{13.58}$$

Similar approximations for the potential are used in the spawning and coherent coupled states (CCS) GWP methods to take advantage of analytic expressions in forming the required integrals. In spawning, under the name of *saddle-point approximation*, only the zero-order energy is used [29]. More recently the bra-ket averaged Taylor (BAT) approximation has been introduced in conjunction with the CCS method [30]. This includes the zero- and first-order terms, and averages over expansions at the centers of the bra and ket GWPs, rather than second-order expansion around the centre of the ket as used in vMCG.

Let us now come back to the CX-formalism, discussed in subsection 13.2.2. The extent of the zeroth-order Hamiltonian is determined by the representability of the Gaussian moments by $\mathbf{S}^{(0\nu)}$. From Eq. (13.17), we easily get

$$M_{jk}^{(0)} = S_{jk}^{(0\eta)} \;;\; M_{jk}^{(1)} = S_{jk}^{(0\xi)} \;;\; M_{jk}^{(2)} = S_{jk}^{(0\zeta)} \,, \tag{13.59}$$

where the constraint on η_k in terms of the other two parameters is not accounted for at this stage. In addition, $M_{jk}^{(0)} = S_{jk}$, which effectively removes any explicit involvement of η_k.

For a frozen GWP, only zeroth- and first-order moments are available from $S_{jk}^{(0\eta)}$ and $S_{jk}^{(0\xi)}$ (getting second-order moments from $S_{jk}^{(0\zeta)}$ requires ζ_k to be varied with time). The corresponding zeroth-order Hamiltonian matrix elements can be written as

$$H_{jl}^0 = S_{jk}^{(0)} X_k^{(0)} + S_{jk}^{(0\xi)} X_k^{(1)} \,, \tag{13.60}$$

Thus, the zeroth-order part of the EOMs for $\xi_k = \xi_k^0 + \xi_k^{\mathrm{R}}$ are of the form

$$\mathrm{i}\hbar \frac{\partial \xi_k^0}{\partial t} = \mathrm{i}\hbar \frac{1}{2\sigma^2}\frac{\partial q_k}{\partial t} - \frac{\partial p_k}{\partial t} = X_k^{(1)} \tag{13.61}$$

$$= \mathrm{i}\hbar \frac{1}{2\sigma^2}\frac{p_k}{\mu_k} + V_k' - 2X_k^{(2)} q_k \,, \tag{13.62}$$

while the classical Newton EOMs for $\frac{\partial q_k}{\partial t}$ and $\frac{\partial p_k}{\partial t}$ have the following form,

$$\frac{\partial q_k}{\partial t} = \frac{p_k}{\mu} \,, \tag{13.63a}$$

$$\frac{\partial p_k}{\partial t} = -V_k' \,. \tag{13.63b}$$

In order to separate the EOMs into classical and non-classical terms, we can choose to put the extra term (a sort of width-induced, "local quantum force"), $-2X_k^{(2)}q_k$, into $\mathbf{Y}^{\mathrm{R}}$. All "quantum" contributions are thus kept in $\mathbf{Y}^{\mathrm{R}}$, and ignoring them means that each GWP will follow a classical trajectory. This slightly modified CX-formalism connects the variational GWPs of vMCG with the classical basis functions of the spawning or *coupled coherent states* (CCS) approaches. A detailed comparison showing the relationship between vMCG and CSS is given in [31].

Finally, it should also be mentioned that for coherent states in a harmonic well, the extra term cancels for the appropriate time-dependent width. Indeed, in Eq. (13.57c), if $\sigma = \sqrt{\frac{\hbar}{2\mu_k\omega_k}}$ with $\omega_k = \sqrt{\frac{V_k''}{\mu_k}}$ (see subsection 13.2.1), then $X_k^{(2)} = 0$. This is consistent with the common knowledge that thawed GWPs in a harmonic well move classically, so that in our notation $\mathrm{i}\hbar\frac{\partial\mathbf{\Lambda}}{\partial t} = \mathbf{X}$ is the complete solution (further details can be found in Refs. [7, 14, 18]). This also implies the decoupling of the parameter-EOMs among all GWPs, which only stay coupled via the expansion coefficients. In contrast, the frozen-width case derived above is, surprisingly, more complicated in terms of the separation of the classical and non-classical contributions [18].

13.3 Example Calculations

There are two classes of system that must be treated with the TDSE rather than classical mechanics for correct results. The first class of molecules for which quantum effects are key are those in which tunneling is significant. This is particularly important for reactions in which a proton transfer is the rate limiting step. In the second class are molecules that possess a manifold of electronic states coupled by vibronic coupling. In many cases, this *non-adiabatic coupling* between the electronic and nuclear motion results in a conical intersection being present between potential-energy surfaces which can provide pathways for ultrafast crossing between electronic states (See Chapter 1). Below we will demonstrate that vMCG can indeed treat both of these types of system.

In addition to capturing the quantum features of a system, a useful quantum-dynamics method has also to show stability and good convergence. Scaling with system size is also a key feature. As the examples show, vMCG has excellent convergence and is an exact numerical method if the integrals are evaluated exactly.

Scaling of the method is usually dominated by the size of the **C**-matrix in the GWP parameter-EOM, Eq. (13.38), that needs to be inverted at each step. For the standard frozen GWP formulation, this matrix scales as $(n \times f)^2$ where n is the number of functions and f the number of degrees of freedom. Thus this can become very large. It is, however, possible to reduce this scaling by using an MCTDH formulation with sets of low-dimensional GWPS in a direct-product expansion of the wave function. The properties of this formulation in terms of accuracy is the same as standard vMCG and we will not look further at this possibility here.

Nothing, however, can be said in general about the number of functions needed, as this is highly dependent on the system studied. As with the MCTDH method the number of basis functions required depends on the strength of correlation among degrees of freedom. If many modes are strongly coupled more functions are required to describe the complicated dynamics than in a system with many lightly coupled modes that effectively undergo harmonic motion: in the limit of uncoupled harmonic oscillators only one GWP is required, irrespective of system size. The number of functions required also depends on the property calculated and the accuracy desired.

Due to the need for the inversion of the **C**-matrix, and also the inversion of the GWP overlap matrix in the coefficient-EOMs, Eq. (13.36), the vMCG EOMs can be numerically unstable. While

the results of a calculation does not depend on the initial positions of the basis functions, in practice this plays a role in the stability. The main choice is whether to distribute initially unpopulated GWPs in coordinate space or momentum space, or randomly in phase space. It appears that distribution in momentum space is the best choice, as all GWPs are then guaranteed to have sensible initial potential energy.

13.4 Tunneling Dynamics: Salicylaldimine

Salicylaldimine is a molecule that can undergo proton transfer between an oxygen and a nitrogen atom attached to an aromatic ring. This makes it a good test system, as the rigidity means that the proton transfer is only minimally coupled to the other vibrations. The proton transfer can be described by two vibrational modes corresponding to the transfer mode, ν_1, and the in-plane proton vibration, ν_{36}. These are shown in Figure 13.1(a). A model potential energy in terms of these modes

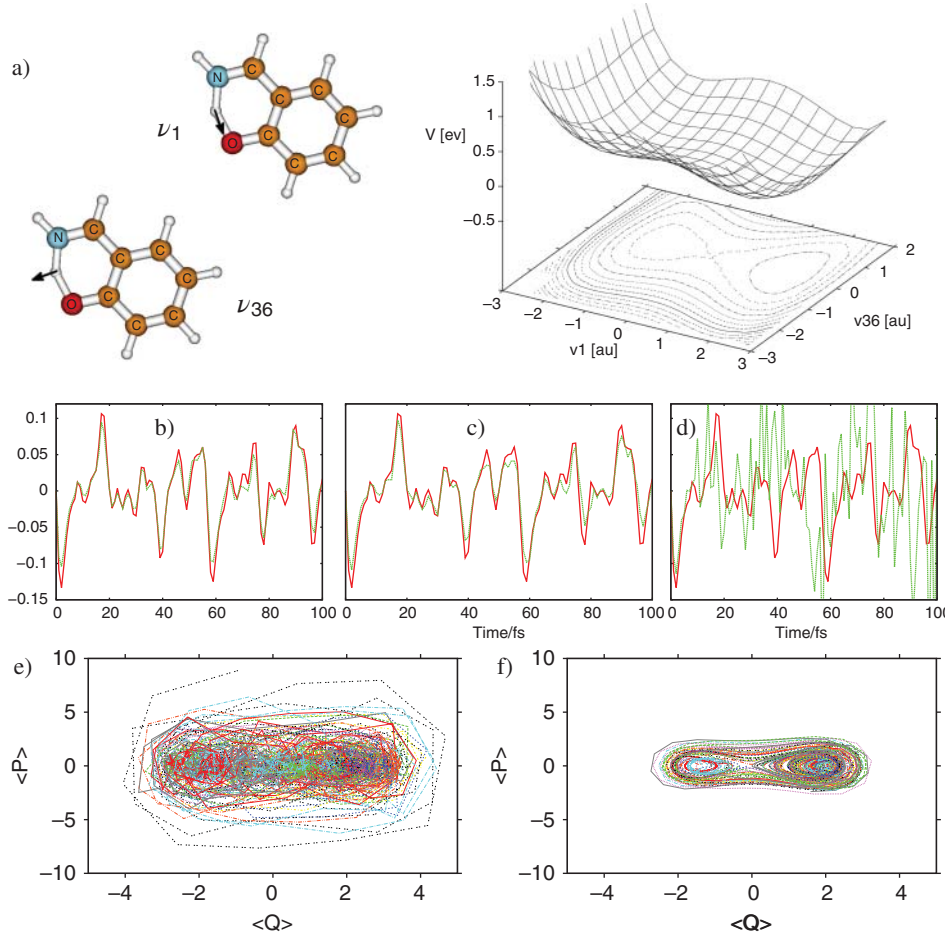


Figure 13.1 (a) Salicylaldimine proton transfer modes and 2D potential-energy surface. (b)–(d) Flux through the barrier for a tunneling dominated reaction. In red is the exact solution and green lines are for (b) 16 vMCG GWPS and fourth order integrals (c) 32 vMCG narrow GWPs and 2nd order integrals (d) 64 classical GWPs and second order integrals. (e), (f) phase space plots for mode ν_1 using (e) 32 vMCG GWPs and (f) 32 classical GWPs.

has been fitted to HF/6-31G* calculations [32] and can be represented by the following expression (in terms of frequency-mass-weighted, dimensionless, normal coordinates),

$$\hat{H} = \sum_{\kappa=1,36} \frac{\hbar\omega_\kappa}{2}\left(\frac{\partial^2}{\partial q_\kappa^2} + q_\kappa^2\right) + \sum_{n=1}^{4} A_n q_1^n + B_{11} q_1 q_{36} + B_{22} q_1^2 q_{36}^2 + B_{31} q_1^3 q_{36} + B_{13} q_1 q_{36}^3 \,. \tag{13.64}$$

The potential is an asymmetric double well and is also shown in Figure 13.1(a).

The flux for proton transfer from oxygen to nitrogen is also shown in Figure 13.1, calculated using MCTDH and vMCG after starting with a small bond compression, but with an initial energy below the barrier height so that transfer is dominated by tunneling. In Figure 13.1(b) the GWP basis functions have a width corresponding to the ground-state vibrational wave packet. These are 0.57 along ν_1 and 0.7071 along ν_{36}. Integrals were calculated to fourth order, i.e., the calculation is complete, and the initial wave packet was composed of a single GWP, i.e., only one GWP is populated. It can be seen that the calculation is converged with 16 GWPs.

If the integrals are performed using an LHA to the potential energy, it is necessary to use narrower GWPs to achieve convergence due to the error in the integrals. Fitting of the GWPs to the initial wave packet is now required to use the same one as before. The flux is shown in Figure 13.1(c), and with widths of 0.25 along ν_1 and 0.65 along ν_{36} is converged with 32 GWPs. Figure 13.1(d) shows the same calculation but with 64 GWPs that follow classical trajectories rather than the coupled vMCG equations of motion. It is clear that the quality of the result is not as good.

To understand why vMCG converges better than a classical GWP basis, phase-space plots for the centers of the basis functions used in vMCG and classical GWPs calculations, both with 32 GWPs starting with the same distribution, are shown in Figures 13.1e,f. The classical GWPs keep to the periodic orbits defined by their initial conditions, whereas the vMCG GWPs spread out to cover phase space better.

13.5 Non-Adiabatic Dynamics: The Butatriene Cation

A classic example of a non-adiabatic system is the butatriene radical cation. The lowest two bands of the photoelectron spectrum are joined by a well-structured band which has been shown to be due to a conical intersection between the states [33]. The spectrum can be well reproduced using a simple *vibronic model* Hamiltonian. In this, a diabatic basis set is assumed, and the Hamiltonian operator is written in a matrix form with potential-energy surfaces and smooth couplings written as Taylor expansions around the Franck–Condon point. This is a standard model for describing a manifold of coupled states and has had much success in describing the short-time dynamics of photo-excited molecules [34, 35]. For the butatriene cation the Hamiltonian can be written (again, in terms of frequency-mass-weighted, dimensionless, normal coordinates)

$$\hat{\mathbf{H}} = \sum_i \frac{\hbar\omega_i}{2}\left(-\frac{\partial^2}{\partial Q_i^2} + Q_i^2\right)\begin{pmatrix}1&0\\0&1\end{pmatrix} + \begin{pmatrix}\epsilon_1&0\\0&\epsilon_2\end{pmatrix} + \sum_{i\in A_g}^{4}\begin{pmatrix}\kappa_i^{(1)}&0\\0&\kappa_i^{(2)}\end{pmatrix} Q_i + \begin{pmatrix}0&\lambda\\\lambda&0\end{pmatrix} Q_{A_u} + \ldots \tag{13.65}$$

Table 13.1 shows the results from four sets of calculations to demonstrate the convergence and scaling properties of vMCG for this model. The four calculations use different sets of modes. The

Table 13.1 Computer resources required for calculating the autocorrelation function of varying size models of the butatriene cation to 100 fs using MCTDH and vMCG convergence on a Linux cluster.

	MCTDH			vMCG		
Model	Config	Time (min)	Mem (MB)	GWPs	Time (min)	Mem (MB)
2D	2100	0.5	1.4	40	6	14
5D	9498	1	4.5	60	93	42
12D	259436	24	151	80	140	1371
18D	466720	63	185	100	399	15449

2D model includes just the ν_5 and ν_{14} vibrations that are modes with the strongest linear coupling responsible for the conical intersection. The 5D model includes the remaining tuning modes with A_g symmetry. The 12D model includes a further seven vibrations that have the most significant second-order terms in the Hamiltonian, while the 18D model includes all modes and are thus complete calculations of the model to second order.

The results show that vMCG requires more computational effort than grid-based MCTDH. In particular the scaling of computer time with system size is poor. This is mainly due to the need for small step sizes and is an area that requires work. One of the main advantages of the method, however, is to enable direct-dynamics calculations in which the potential energies are computed on-the-fly rather than requiring pre-computed functions. The GWP basis functions combined with an LHA of the potential energies make it straightforward to compute the integrals required using the energy, gradient, and Hessian calculated using an electronic-structure code; the Quantics program has been interfaced to a number of such codes to allow this to happen. Details of the implementation are to be found in Ref. [23] but the main novelty is the use of a "database" in which the quantum-chemistry results are stored. These are not calculated at every step along the propagation but only when new points of data are required. Interpolation is then used to provide the global surfaces. This has the advantage of not only reducing the number of (expensive) quantum-chemistry calculations required but also providing potential-energy surfaces that can be used for analysis and further refinement.

A drawback of using an LHA is that calculations of non-adiabatic systems must be run in the diabatic picture to ensure continuity when crossing the seam. The question then is how can these diabatic surfaces be defined when the quantum-chemistry calculations provide adiabatic quantities? The simplest general option is to use a "propagation diabatization" scheme, which makes use of the relationship

$$\nabla \mathbf{U} = -\mathbf{F}\mathbf{U} \,, \tag{13.66}$$

where $\mathbf{F}$ is the (first-order) non-adiabatic coupling matrix and $\mathbf{U}$ is the diabatic-to-adiabatic basis transformation matrix that is propagated along with the basis functions [36]. Figure 13.2 shows the surfaces produced during a DD-vMCG simulation of the butatriene cation including the first and second states. The conical intersection between the states can be seen in the adiabatic surfaces, and the diabatic surfaces are smooth intersecting sheets, as desired. It should be kept in mind that Eq. (13.66) is only exact for a full manifold of states. Truncating the manifold, as is done in all practical calculations, may result in a path dependency of the calculated transformation matrices, and further work is needed to see how important this may be.

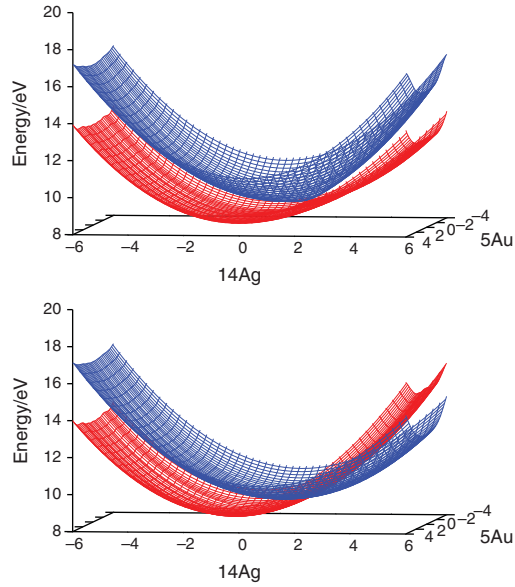
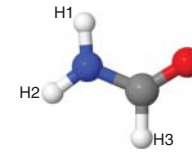


Figure 13.2 Potential-energy surfaces of the butatriene cation calculated at the SA-CAS(5,6)/3-21G level of theory from DD-vMCG simulations. (a) Adiabatic surfaces; the red surface corresponds to the D_0 state, the blue to D_1. (b) Diabatic surfaces; the red surface corresponds to the $\tilde{X}$ state, the blue to $\tilde{A}$.

13.6 Direct Non-Adiabatic Dynamics: Formamide

The potential of DD-vMCG calculations using the propagation diabatization for studying non-adiabatic systems is shown by the results from preliminary calculations on formamide [37]. This is the smallest molecule containing C, N, and O. As a building block of DNA bases, it is thought to have been a key part of the initiation of life after irradiation to photochemically form formimidic acid. In Table 13.2 the energy and characterization are given for each of the first seven

Table 13.2 Formamide energies, oscillator strengths, and coefficients of main configurations (values > 0.15) from an SA-8-CAS(10,8)/6-31G* calculation. OLP: oxygen lone pair; NH+NH: orbital is on NH_2 with a node between the NH bonds

	Energy (eV)	Oscillator strength (a.u.)	Main configurations
			
S_1	5.61	0.0008	$0.69(\text{OLP}–\pi^*)$
S_2	8.02	0.0004	$0.66(\pi - \text{NH+NH})$
S_3	8.16	0.0225	$0.54(\text{OLP} - \text{NH+NH}) + 0.34(\pi - \pi^*)$
S_4	9.12	0.0000	$0.66(\pi - \text{NH+NH})$
S_5	10.03	0.0710	$0.63(\text{OLP} - \text{NH+NH})$
S_6	10.57	0.7258	$0.44(\pi - \pi^*) + 0.37\,(\text{OLP} - \text{NH+NH})$
S_7	11.45	0.0013	$0.55(\pi_1 - \text{NH+NH}) + 0.36(\pi - \pi^*)$

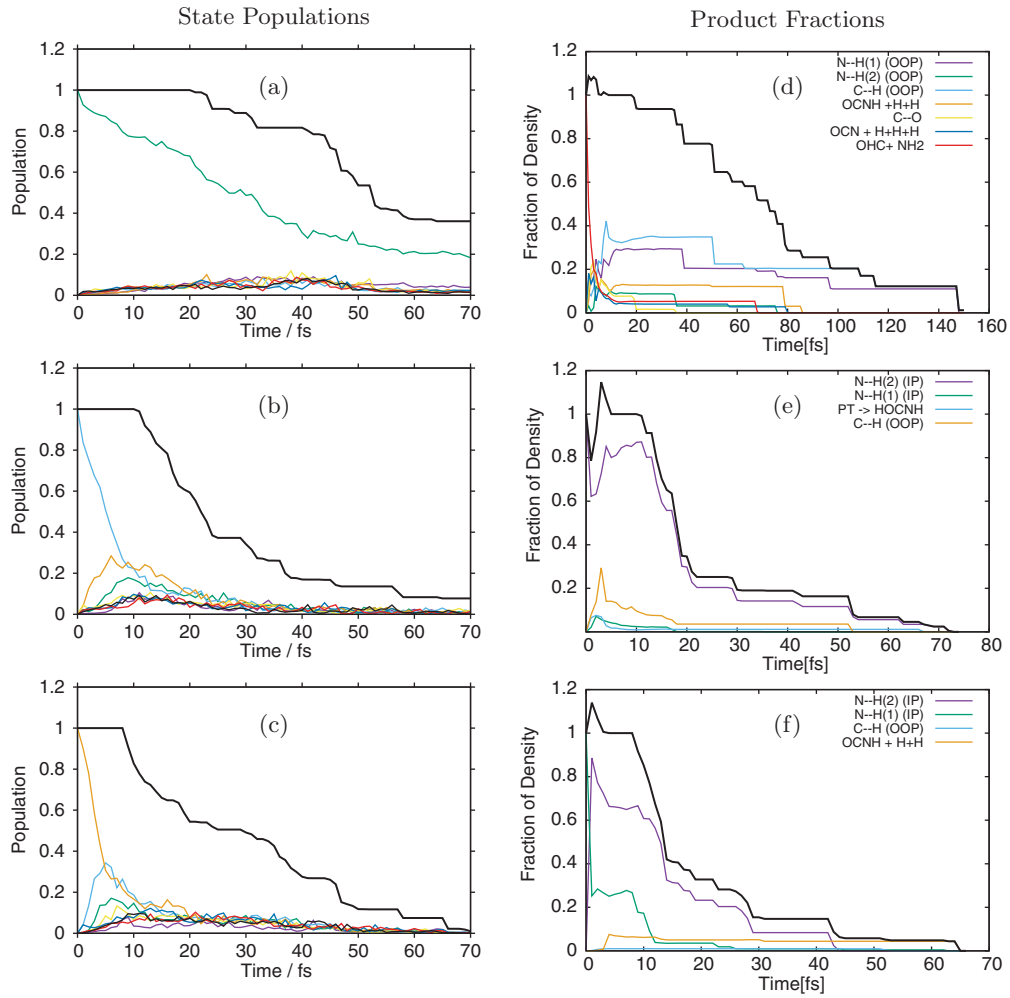


Figure 13.3 Results from DD-vMCG simulations of formamide with SA-8-CAS(10,8)/6-31G* potential-energy surfaces. Left column – Diabatic state populations starting in different states: (a) S_1; (b) S_2; (c) S_3. S_0: purple; S_1: green; S_2: light blue; S_3: orange; S_4: yellow; S_5: dark blue; S_6: red; S_7: black. Total density (squared norm): thick black. Right column – The fraction of density going into different product channels starting in different states: (d) S_1; (e) S_2; (f) S_3. Each line represents the different products defined either by the bond that breaks or the products formed. IP and OOP signify if the dissociation occurs in- or out-of-plane.

excited singlet states calculated at the SA-CAS(10,8)/6-31G* level of theory. These calculations demonstrate the sensitivity of excited-state calculations to the choice of number of states included in addition to the size of the CAS space and basis set. The S_3 state, which has some $\pi\pi^*$ character, is the first state with significant oscillator strength.

In Figure 13.3 the results from DD-vMCG simulations starting with vertical excitations to the lowest three excited states are shown. In addition to the state populations an analysis of the products formed is shown. Each simulation used 48 GWPs. A clear difference in behavior of the molecule after excitation to S_1 is seen compared to that after excitation to S_2 or S_3. Looking at the state populations, both S_2 and S_3 (Figure 13.3(b,c)) undergo very fast decay, seen by the decay of the total

population (black line), accompanied by population transfer predominantly to the neighboring states. The decay is due to *complex absorbing potentials* (CAPs) placed along each axis to remove parts of the wave packet that dissociate. In both cases the decay is completed by 70 fs. In contrast, after excitation to S_1 the dissociation is slower and seems to be equilibrating with 20 % of the wave packet left as formamide in the S_1 state (Figure 13.3). Minimal population transfer to the other states is seen.

To analyze the products formed, the trajectories at the centre of the basis functions can be examined. Each trajectory can then be weighted by the *gross Gaussian population* (GGP) which takes the overlap between the basis functions and apportions the density equally between the overlapping functions. The GGP for a function i is defined as

$$\mathrm{GGP}_i = \Re \sum_j A_i^* S_{ij} A_j \ . \tag{13.67}$$

The fraction of a wave packet ending in different products is also shown in Figure 13.3. Again the behavior is different for S_1 (Figure 13.3(d)) and S_2 or S_3 (Figures 13.3(e,f)). In the latter cases the majority of the product is formed by a direct hydrogen loss from the nitrogen via an in-plane motion. In contrast, from S_1 a number of fragments are formed and hydrogen loss from the nitrogen is out-of-plane.

Cuts through the potential-energy surfaces resulting from the DD-vMCG calculations in both the diabatic and adiabatic representations are shown in Figure 13.4. It can be seen that, again, the diabatic curves possess smooth crossings, while the adiabatic ones contain cusps at crossing points. An important feature that can be seen in the cuts along the N–H stretch in Figure 13.4(a,b) is that

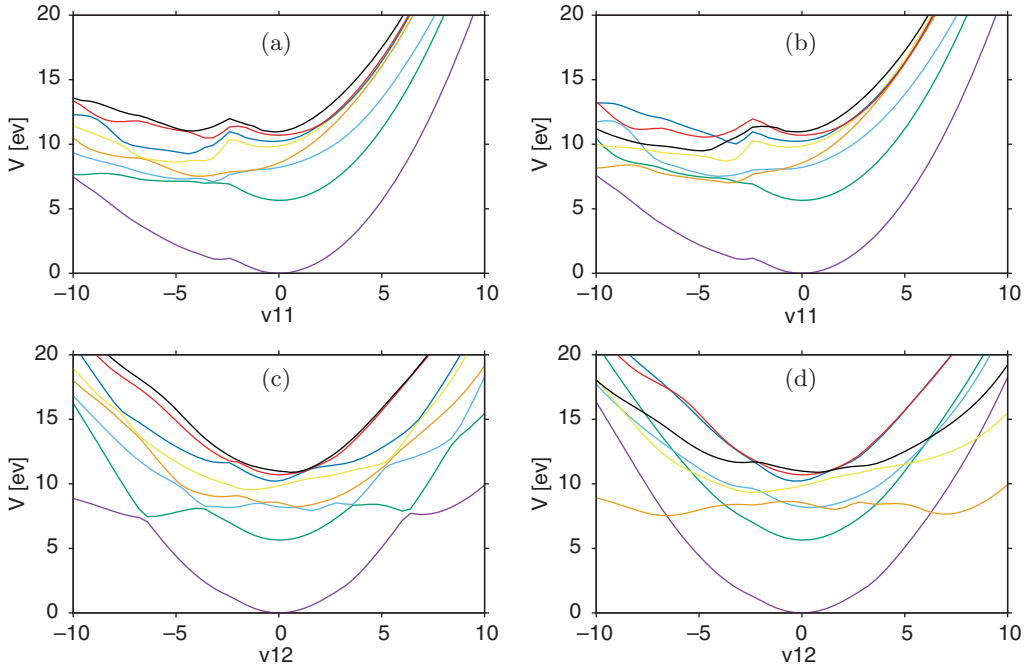


Figure 13.4 Cuts through the SA-8-CAS(10,8)/6-31G* potential-energy surfaces of formamide from DD-vMCG simulations. (a) ν_{11} (N-H_2 symmetric stretch) adiabatic; (b) ν_{11} diabatic; (c) ν_{12} (N-H_2 antisymmetric stretch) adiabatic; (d) ν_{12} diabatic.

the electronic-structure level of theory is clearly breaking down along the dissociation, resulting in a discontinuity.

13.7 Summary

The vMCG method is a flexible and potentially accurate way of treating an evolving wave packet. The GWP basis functions are directly coupled using variational EOMs, both for expansion coefficients and Gaussian parameters, which leads to fast convergence. It also means that quantum effects such as tunneling or non-adiabatic crossing are correctly treated. The variational nature of the EOMs also implies that the best result is obtained irrespective of the initial placement of the GWPs, making it, in principle, a black-box method.

The method can be used either with traditional analytical potential-energy functions or in a direct-dynamics mode. However, the present implementation is somewhat resource hungry, and the numerical problems prevent really accurate calculations. Future work will be required to solve these issues. It is, however, a perfect method for initial explorations of a system to test the stability and accuracy of electronic-structure calculations before doing large-scale potential-energy fitting. It can also be used to calculate points relevant for dynamics. In short, it is a method with great potential in the field of quantum dynamics.

13.8 Practical Implementation

The vMCG and DD-vMCG methods are available in the Quantics Package [25]. This is a general suite of programs able to solve the time-dependent Schrödinger Equation using a number of different algorithms. In addition to the GWP methods described in this chapter, the package contains the full set of grid-based and MCTDH methods described in Chapter 12. There are a range of analysis programs to plot typical quantities such as spectra and state populations. Input of both the system and operator are via text files. Quantics is interfaced to a number of electronic structure codes (Gaussian, Molpro, QChem and Molcas) to enable direct dynamics calculations. The code is available on request to the authors.

Acknowledgments

We would like to acknowledge all the people who have been involved in aspects of the development and implementation of the vMCG and DD-vMCG methods. These include Charlotte Allan, Mike Bearpark, Terry Frankcombe, Kousik Giri, Werner Koch, João Malhado, David Mendive-Tapia, Iakov Polyak, Gareth Richings, Mike Robb, Sarah Römer, Matthias Ruckenbauer, and Eryn Spinlove all of whom have worked on aspects of the theory and / or implementation.

References

1 Heller, E.J. (1975). Time-dependent approach to semiclassical dynamics. *J. Chem. Phys.* 62: 1544–1555.

2 Heller, E.J. (1981). Frozen Gaussians: a very simple semiclassical approximation. *J. Chem. Phys.* 75: 2923–2931.

3 Lee, S.Y. and Heller, E.J. (1982). Exact time-dependent wave packet propagation: application to the photodissociation of methyl iodide. *J. Chem. Phys.* 76: 3035–3044.

4 Sepúlveda, M., Tomsovic, S., and Heller, E. (1992). Semiclassical propagation: how long can it last? *Phys. Rev. Lett.* 69: 402–405.

5 Heller, E.J. (1991). Cellular dynamics: a new semiclassical approach to time-dependent quantum mechanics. *J. Chem. Phys.* 94: 2723–2729.

6 Huber, D. and Heller, E.J. (1987). Generalized Gaussian wave packet dynamics. *J. Chem. Phys.* 87: 5302–5311.

7 Sawada, S.-I., Heather, R., Jackson, B., and Metiu, H. (1985). A strategy for time dependent quantum mechanical calculations using a Gaussian wave packet representation of the wave function. *J. Chem. Phys.* 83: 3009–3027.

8 Sawada, S.I. and Methiu, H. (1986). A Gaussian wave packet method for studying time dependent quantum mechanics in a curve-crossing system: low energy motion, tunneling and thermal dissipation. *J. Chem. Phys.* 84: 6293–6311.

9 Sawada, S. and Methiu, H. (1986). A multiple trajectory theory for curve crossing problems obtained by using a Gaussian wave packet representation of the nuclear motion. *J. Chem. Phys.* 84: 227–238.

10 Martínez, T.J., Ben-Nun, M., and Levine, R.D. (1996). Multi-electronic-state molecular dynamics: a wave function approach with applications. *J. Phys. Chem.* 100: 7884–7895.

11 Ben-Nun, M. and Martínez, T.J. (1998). Nonadiabatic molecular dynamics: validation of the multiple spawning method for a multidimensional problem. *J. Chem. Phys.* 108: 7244–7257.

12 Martínez, T.J. and Levine, R.D. (1996). First-principles molecular dynamics on multiple electronic states: a case study of NaI. *J. Phys. Chem.* 105: 6334–6341.

13 Ben-Nun, M. and Martínez, T.J. (2002). *Ab Initio* Quantum Molecular Dynamics. *Adv. Chem. Phys.* 121: 439–512.

14 Burghardt, I., Meyer, H.D., and Cederbaum, L.S. (1999). Approaches to the approximate treatment of complex molecular systems by the multiconfiguration time-dependent Hartree method. *J. Chem. Phys.* 111: 2927–2938.

15 Worth, G.A. and Burghardt, I. (2003). Full quantum mechanical molecular dynamics using Gaussian wave packets. *Chem. Phys. Lett.* 368: 502–508.

16 Worth, G.A., Robb, M.A., and Burghardt, I. (2004). A novel algorithm for non-adiabatic direct dynamics using variational Gaussian wave packets. *Faraday Discuss.* 127: 307–323.

17 Lasorne, B., Robb, M.A., and Worth, G.A. (2007). Direct quantum dynamics using variational multi-configuration Gaussian wave packets. Implementation details and test case. *PCCP* 9: 3210–3227.

18 Burghardt, I., Giri, K., and Worth, G.A. (2008). Multimode quantum dynamics using Gaussian wave packets: the Gaussian-based multiconfiguration time-dependent Hartree (G-MCTDH) method applied to the absorption spectrum of pyrazine. *J. Chem. Phys.* 129: 174104–174114.

19 Worth, G.A., Robb, M.A., and Lasorne, B.L. (2008). Solving the time-dependent Schrödinger equation for nuclear motion in one step: direct dynamics of non-adiabatic systems. *Mol. Phys.* 106: 2077–2091.

20 Lasorne, B. and Worth, G.A. (2009). Direct dynamics of quantum nuclei. In: *MCTDH: Basic Theory, Extensions and Applications to Multidimensional Quantum Dynamics* (eds. H.D. Meyer, F. Gatti and G.A. Worth), 113–129. Weinheim, Germany: VCH.

21 Allan, C.S.M., Lasorne, B., Worth, G.A., and Robb, M.A. (2010). A straightforward method of analysis for direct quantum dynamics: application to the photochemistry of a model cyanine. *J. Phys. Chem. A* 114: 8713–8729.

22 Mendive-Tapia, D., Lasorne, B., Worth, G.A., et al. (2012). Towards converging non-adiabatic direct dynamics calculations using frozen-width variational Gaussian product basis functions. *J. Chem. Phys.* 548: 22A548–22A510.

23 Richings, G.W., Polyak, I., Spinlove, K.E., et al. (2015). Quantum dynamics simulations using Gaussian wave packets: the vMCG method. *Int. Rev. Phys. Chem.* 34: 269–308.

24 Richings, G.W. and Worth, G.A. (2015). A practical diabatisation scheme for use with the direct-dynamics variational multi-configuration Gaussian method. *J. Phys. Chem. A* 119: 12457–12470.

25 Worth, G.A. (2020). Quantics: a general purpose package for quantum molecular dynamics simulations. *Comput. Phys. Commun.* 248: 107040.

26 Dirac, P.A.M. (1930). Note on exchange phenomena in the Thomas atom. *Proc. Camb. Philos. Soc.* 26: 376–385.

27 Frenkel, J. (1934). *Wave Mechanics*. Oxford, U.K: Clarendon Press.

28 Römer, S. and Burghardt, I. (2013). Towards a variational formulation of mixed quantum-classical molecular dynamics. *Mol. Phys.* 111: 3618–3624.

29 Martínez, T.J., Ben-Nun, M., and Levine, R.D. (1997). Molecular collision dynamics on several electronic states. *J. Phys. Chem. A* 101: 6389–6402.

30 Makhov, D.V., Glover, W.J., Martínez, T.J., and Shalashilin, D.V. (2014). Ab initio multiple cloning algorithm for quantum non-adiabatic molecular dynamics. *J. Chem. Phys.* 141: 54110–54111.

31 Shalashilin, D.V. and Burghardt, I. (2008). Gaussian-based techniques for quantum propagation from the time-dependent variational principle: formulation in terms of trajectories of coupled classical and quantum variables. *J. Chem. Phys.* 129: 84104–84109.

32 Polyak, I., Allan, C.S.M., and Worth, G.A. (2015). A complete description of tunnelling using direct quantum dynamics simulation: salicylaldimine proton transfer. *J. Chem. Phys.* 143: 84121.

33 Cederbaum, L.S. et al. (1977). Strong vibronic coupling effects in ionization spectra: the "mystery band" of butatriene. *Chem. Phys.* 26: 169–177.

34 Köppel, H. and Meyer, H.D. (1984). Novel aspects of ultrafast non-radiative processes. *Chem. Phys. Lett.* 107: 149–154.

35 Worth, G.A., Meyer, H.-D., Köppel, H., et al. (2008). Using the MCTDH wave packet propagation method to describe multimode non-adiabatic dynamics. *Int. Rev. Phys. Chem.* 27: 569–606.

36 Richings, G.W. and Worth, G.A. (2017). Multi-state non-adiabatic direct-dynamics on propagated diabatic potential energy surfaces. *Chem. Phys. Lett.* 683: 606–612.

37 Spinlove, K.E., Richings, G.W., Robb, M.A., Worth, G.A. (2018). Curve crossing in a manifold of coupled electronic states: direct quantum dynamics simulations of formamide. *Faraday Discuss.* 212: 191–215.

14

Full and *Ab Initio* Multiple Spawning

Basile F. E. Curchod

Department of Chemistry, Durham University, Durham DH1 3LE, United Kingdom

Abstract

This Chapter describes the method for non-adiabatic quantum molecular dynamics called Full Multiple Spawning. The Full Multiple Spawning framework proposes to portray nuclear wave functions by linear combinations of classically-traveling multi-dimensional Gaussian functions, called trajectory basis functions. The number of trajectory basis functions can be adapted when needed during the excited-state dynamics through a spawning algorithm, and all basis functions are coupled together.

Hence, Full Multiple Spawning is a formally exact method for non-adiabatic dynamics in the limit of a large number of basis functions. Full Multiple Spawning can be extended to the description of light/matter interaction or the inclusion of spin-orbit coupling effects. Two controlled approximations can be performed on the Full Multiple Spawning equations and lead to the Ab Initio Multiple Spawning technique, which allows for on-the-fly non-adiabatic quantum dynamics of medium-size molecules. In addition to describing the formalism and algorithms of the Full- and Ab Initio Multiple Spawning, this Chapter presents a dissection of a typical Ab Initio Multiple Spawning dynamics run, as well as different successful applications of this technique.

14.1 Introduction

Simulating the excited-state dynamics of a molecular system requires going beyond the Born–Oppenheimer approximation, accounting for certain critical nuclear quantum effects, usually neglected in ground-state dynamics. Two main families of methods have emerged to tackle this challenge: (*i*) expressing the time-dependent Schrödinger equation on a grid (Figure 14.1(a), see also Chapter 11) or (*ii*) representing nuclear wave packets by a swarm of trajectories – whether they are quantum, classical, coupled, or uncoupled (Figure 14.1(b)). This chapter focuses on a strategy called *Full Multiple Spawning* (FMS), which aims to offer an *in principle* exact framework to perform *ab initio* non-adiabatic quantum molecular dynamics, that is, simulating the excited-state quantum dynamics of molecules in their full dimensionality and without the need to precompute electronic structure quantities. FMS achieves this goal by combining features from both of the strategies mentioned above, portraying nuclear wave functions as a linear combination of frozen Gaussian functions, which are not static but will evolve in position and momentum

Quantum Chemistry and Dynamics of Excited States: Methods and Applications,
First Edition. Edited by Leticia González and Roland Lindh.

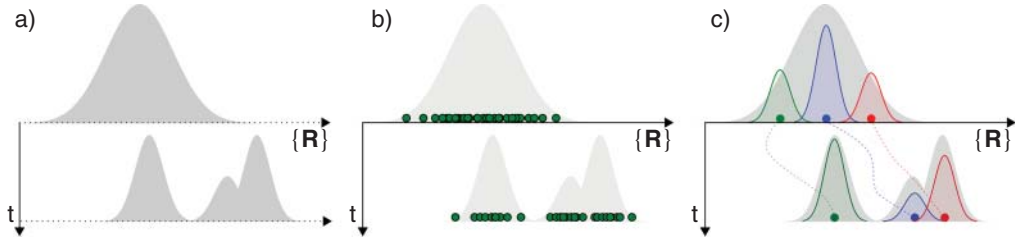


Figure 14.1 Different strategies to picture nuclear wave functions in excited-state dynamics (the horizontal axis represents a given nuclear coordinate, while the vertical axis corresponds to time). (a) The nuclear wave function (grey area) can be expressed on a fixed grid (dots in the $\boldsymbol{R}$ direction) that does not change over time. (b) The probability density of the nuclear wave function can be portrayed by a swarm of quantum or classical trajectories (black-green dots, moving over time). (c) The nuclear wave function can be expanded in a basis of coupled Gaussian functions that can move over time.

space (Figure 14.1(c)) – a sort of moving grid of Gaussian functions. The idea of describing nuclear wave packets utilizing frozen, traveling Gaussian functions is rooted in the seminal work of Heller [23–25]. Since then, different flavors of this strategy have appeared in the context of non-adiabatic dynamics and mostly differ from each other by their choice of equations of motion for the Gaussians.

In this chapter, we begin by discussing the representation of the time-dependent molecular Schrödinger equation in a basis of moving Gaussian functions, as well as the different ways of propagating these Gaussian functions (Section 14.2). We then introduce the FMS framework (Section 14.3), focusing on the so-called spawning algorithm as well as on the critical coupling between Gaussian functions. We present different extensions of FMS (Section 14.4), which aim to include additional physical processes in the FMS framework, such as spin–orbit coupling or the effect of an external time-dependent electromagnetic field. We then discuss the central approximations applied to the FMS equations to generate the *Ab Initio* Multiple Spawning (AIMS) technique (Section 14.5), allowing for on-the-fly non-adiabatic dynamics of molecular systems. We comment on the quality of AIMS approximations and put them in perspective with those of surface hopping for the treatment of non-adiabatic events. Finally we describe how an AIMS simulation is performed, provide a detailed example of an AIMS run (Section 14.6), and present a series of successful applications of the method to molecular systems (Section 14.7).

14.2 Time-Dependent Molecular Schrödinger Equation in a Gaussian Basis

14.2.1 Central Equations of Motion

In a non-relativistic framework, the time-dependent Schrödinger equation describes the dynamics of electrons and nuclei in a molecular system and is written as

$$\hat{H}(\boldsymbol{r},\boldsymbol{R})\Psi(\boldsymbol{r},\boldsymbol{R},t) = i\frac{\partial}{\partial t}\Psi(\boldsymbol{r},\boldsymbol{R},t)\ , \tag{14.1}$$

where $\Psi(\boldsymbol{r},\boldsymbol{R},t)$ is the total molecular wave function, with $\boldsymbol{r} = (\boldsymbol{r}_1,\ldots,\boldsymbol{r}_i,\ldots,\boldsymbol{r}_{n_{el}})$, $\boldsymbol{R} = (\boldsymbol{R}_1,\ldots,\boldsymbol{R}_I,\ldots,\boldsymbol{R}_{N_n})$ (n_{el} being the number of electrons and N_n the number of nuclei). The molecular Hamiltonian $\hat{H}(\boldsymbol{r},\boldsymbol{R})$ contains the kinetic energy operator for both the electrons and the

nuclei, as well as all the interaction potentials between electrons and nuclei (i,j label the electrons, I,J the nuclei; atomic units are used throughout this chapter[1]):

$$\begin{aligned}\hat{H}(\boldsymbol{r},\boldsymbol{R}) &= -\sum_{I}^{N_n}\frac{1}{2M_I}\nabla_I^2 - \sum_{i}^{n_{el}}\frac{1}{2}\nabla_i^2 + \sum_{i<j}\frac{1}{|\boldsymbol{r}_i-\boldsymbol{r}_j|} - \sum_{I,i}\frac{Z_I}{|\boldsymbol{R}_I-\boldsymbol{r}_i|} + \sum_{I<J}\frac{Z_I Z_J}{|\boldsymbol{R}_I-\boldsymbol{R}_J|} \\ &= -\sum_{I}^{N_n}\frac{1}{2M_I}\nabla_I^2 + \hat{\mathcal{H}}_{el}(\boldsymbol{r},\boldsymbol{R}) = \hat{T}_n(\boldsymbol{R}) + \hat{\mathcal{H}}_{el}(\boldsymbol{r},\boldsymbol{R})\ . \end{aligned} \tag{14.2}$$

$\hat{\mathcal{H}}_{el}(\boldsymbol{r},\boldsymbol{R})$ is called the electronic (or Born–Oppenheimer) Hamiltonian. For a fixed set of nuclear coordinates, one can write an eigenvalue equation for this Hamiltonian,

$$\hat{\mathcal{H}}_{el}(\boldsymbol{r},\boldsymbol{R})\Phi_\nu(\boldsymbol{r};\boldsymbol{R}) = E_\nu^{el}(\boldsymbol{R})\Phi_\nu(\boldsymbol{r};\boldsymbol{R})\ , \tag{14.3}$$

defining the time-independent electronic Schrödinger equation. The eigenfunctions of Eq. (14.3), $\{\Phi_\nu(\boldsymbol{r};\boldsymbol{R})\}$, are the electronic (adiabatic) wave functions, while the eigenvalues $\{E_\nu^{el}(\boldsymbol{R})\}$ are called electronic energies (ν is a label for the electronic state under consideration). More information about the adiabatic (and diabatic) representation can be found in Chapter 1.

In the so-called Born–Huang (BH) representation [9, 15], the electronic eigenstates defined above are used as an orthonormal basis to express the total molecular wave function,

$$\Psi(\boldsymbol{r},\boldsymbol{R},t) = \sum_{\nu}^{\infty}\Phi_\nu(\boldsymbol{r};\boldsymbol{R})X_\nu(\boldsymbol{R},t)\ , \tag{14.4}$$

where $X_\nu(\boldsymbol{R},t)$ are time-dependent expansion coefficients, usually associated with a time-dependent nuclear wave function in electronic state ν. We note here that, while the BH ansatz is one of the most commonly employed representations for the molecular wave function, others are possible – for example the exact factorization, described in detail in Chapter 17. The BH representation (Eq. (14.4)) is at the heart of our portrayal of photo-physical and photo-chemical processes in molecules, i.e., time-dependent nuclear wave functions evolving on time-independent potential energy surfaces. In the limit where one restricts the BH summation to only one electronic state, the nuclei would be forced to evolve on a single electronic adiabatic eigenstate – the so-called Born–Oppenheimer approximation [83]. Here, we will focus on the more general case where non-adiabatic effects, that is, couplings between nuclear motion and electronic states, are fully accounted for.

In the following, we shall investigate a special strategy to describe the nuclear amplitudes in Eq. (14.4): the use of a basis set composed of moving multi-dimensional Gaussian functions, called *trajectory basis functions* (TBFs),

$$X_\nu(\boldsymbol{R},t) = \sum_{k}^{N_{TBFs}^{\nu}} C_k^{(\nu)}(t)\tilde{\chi}_k^{(\nu)}(\boldsymbol{R};\bar{\boldsymbol{R}}_k^{(\nu)}(t),\bar{\boldsymbol{P}}_k^{(\nu)}(t),\boldsymbol{\alpha}_k^{(\nu)}(t),\bar{\gamma}_k^{(\nu)}(t))\ . \tag{14.5}$$

This equation expresses a given nuclear amplitude $X_\nu(\boldsymbol{R},t)$ as a linear combination of N_{TBFs}^{ν} multi-dimensional Gaussian functions $\tilde{\chi}_k^{(\nu)}(\boldsymbol{R};\bar{\boldsymbol{R}}_k^{(\nu)}(t),\bar{\boldsymbol{P}}_k^{(\nu)}(t),\boldsymbol{\alpha}_k^{(\nu)}(t),\bar{\gamma}_k^{(\nu)}(t))$, each of them associated with a time-dependent complex coefficient $C_k^{(\nu)}(t)$ (k labels a specific TBF, here evolving in electronic state ν). The multi-dimensional Gaussian functions are centered at the position $\bar{\boldsymbol{R}}_k^{(\nu)}(t)$ and momentum $\bar{\boldsymbol{P}}_k^{(\nu)}(t)$, have a time-dependent width matrix $\boldsymbol{\alpha}_k^{(\nu)}(t)$ as well as a phase $\bar{\gamma}_k^{(\nu)}(t)$. It is key here to realize that the Gaussian functions will be moving over time, i.e., their centers

1 In this system of units, $\hbar$, m_e, e, and $\frac{1}{4\pi\epsilon_0}$ are all equal to 1 atomic unit (a.u.) and their respective symbol will be omitted from the equations.

(position and momentum) are time-dependent and not just fixed – as it would be the case for a simple grid (see Figure 14.1).

Each multi-dimensional Gaussian function is described as a product of one-dimensional Gaussian functions,

$$\tilde{\chi}_k^{(\nu)}(\boldsymbol{R};\bar{\boldsymbol{R}}_k^{(\nu)}(t),\bar{\boldsymbol{P}}_k^{(\nu)}(t),\boldsymbol{\alpha}_k^{(\nu)}(t),\bar{\gamma}_k^{(\nu)}(t)) = e^{i\bar{\gamma}_k^{(\nu)}(t)}\prod_{\rho}^{3N_n}\tilde{\chi}_{k\rho}^{(\nu)}(R_\rho;\bar{R}_{k\rho}^{(\nu)}(t),\bar{P}_{k\rho}^{(\nu)}(t),\alpha_{k\rho}^{(\nu)}(t)) \,, \tag{14.6}$$

where ρ runs over all $3N_n$ nuclear coordinates. Each one-dimensional Gaussian function is then expressed as

$$\tilde{\chi}_{k\rho}^{(\nu)}(R_\rho;\bar{R}_{k\rho}^{(\nu)}(t),\bar{P}_{k\rho}^{(\nu)}(t),\alpha_{k\rho}^{(\nu)}(t)) = \left(\frac{2\alpha_{k\rho}}{\pi}\right)^{1/4}\exp[-\alpha_{k\rho}(R_\rho-\bar{R}_{k\rho}^{(\nu)})^2 + i\bar{P}_{k\rho}^{(\nu)}(R_\rho-\bar{R}_{k\rho}^{(\nu)})] \,. \tag{14.7}$$

To summarize, we will attempt in the following to solve the time-dependent molecular Schrödinger equation using a BH representation for the molecular wave function, where each nuclear amplitude will itself be expressed as a linear combination of TBFs[2], that is,

$$\Psi(\boldsymbol{r},\boldsymbol{R},t) = \sum_{\nu}^{\infty}\sum_{k}^{N_{TBFs}^{\nu}} C_k^{(\nu)}(t)\tilde{\chi}_k^{(\nu)}(\boldsymbol{R};\bar{\boldsymbol{R}}_k^{(\nu)}(t),\bar{\boldsymbol{P}}_k^{(\nu)}(t),\boldsymbol{\alpha}_k^{(\nu)}(t),\bar{\gamma}_k^{(\nu)}(t))\Phi_\nu(\boldsymbol{r};\boldsymbol{R}) \,. \tag{14.8}$$

At this stage, a large number of questions should come to mind: how should we evolve the TBFs? How many TBFs do we need? How should we set the parameters of the Gaussian functions? We shall answer these questions progressively as we define the different strategies employing TBFs for non-adiabatic dynamics.

What happens if we (*i*) inject our TBFs version of the BH representation (Eq. (14.8)) into the time-dependent molecular Schrödinger equation (Eq. (14.1)), (*ii*) left multiply the result by

$$[\Phi_\mu(\boldsymbol{r};\boldsymbol{R})\chi_{k'}^{(\mu)}(\boldsymbol{R};\bar{\boldsymbol{R}}_{k'}^{(\mu)}(t),\bar{\boldsymbol{P}}_{k'}^{(\mu)}(t),\boldsymbol{\alpha}_{k'}^{(\mu)}(t),\bar{\gamma}_{k'}^{(\mu)}(t))]^* \,, \tag{14.9}$$

and integrate over both electronic and nuclear coordinates? We obtain a set of equations of motion in a matrix form for all the complex amplitudes,

$$\frac{d\boldsymbol{C}}{dt} = -i\boldsymbol{S}^{-1}[(\boldsymbol{H}-i\dot{\boldsymbol{S}})\boldsymbol{C}] \,. \tag{14.10}$$

This equation, which is in fact the time-dependent molecular Schrödinger equation expressed in a basis of TBFs, is composed of the overlap matrix $\boldsymbol{S}$, the Hamiltonian matrix $\boldsymbol{H}$, and the overlap matrix including the time derivative of the basis functions $\dot{\boldsymbol{S}}$. We can give explicit expressions for each element of these matrices by selecting a TBF k (in electronic state μ) and k' (in electronic state ν):

$$(\boldsymbol{S})_{kk'}^{\mu\nu} = \langle\tilde{\chi}_k^{(\mu)}|\tilde{\chi}_{k'}^{(\nu)}\rangle_{\boldsymbol{R}}\delta_{\mu\nu} \,, \tag{14.11}$$

$$(\boldsymbol{H})_{kk'}^{\mu\nu} = \langle\Phi_\mu\tilde{\chi}_k^{(\mu)}|\hat{H}|\tilde{\chi}_{k'}^{(\nu)}\Phi_\nu\rangle_{\boldsymbol{r},\boldsymbol{R}} \,, \tag{14.12}$$

$$(\dot{\boldsymbol{S}})_{kk'}^{\mu\nu} = \langle\tilde{\chi}_k^{(\mu)}|\frac{\partial}{\partial t}\tilde{\chi}_{k'}^{(\nu)}\rangle_{\boldsymbol{R}}\delta_{\mu\nu} \,. \tag{14.13}$$

We will provide a more detailed description of these matrix elements in Section 14.3, in particular those of the Hamiltonian matrix $\boldsymbol{H}$, focusing on how they are computed, how one can extend them to include new physics in the dynamics, and also, in Section 14.5, how one can approximate them.

It is pivotal to realize that Eq. (14.10) couples all the TBFs together, and that – in the limit of a large number of TBFs – the solution of Eq. (14.10) would tend to a numerically exact solution of the time-dependent Schrödinger equation. However, the idea in the following is to capitalize on the

2 We note here that alternative formulations of the total molecular wave function using TBFs were proposed [29, 42].

dynamics of the TBFs to reduce as much as possible the number of basis functions required for the dynamics, N_{TBFs}, by ensuring that they constitute a proper support for the nuclear wave functions. This goal will be achieved by playing with the time-dependence of the Gaussian parameters, namely $\bar{\boldsymbol{R}}_k^{(\nu)}(t)$, $\bar{\boldsymbol{P}}_k^{(\nu)}(t)$, $\boldsymbol{\alpha}_k^{(\nu)}(t)$, and $\bar{\gamma}_k^{(\nu)}(t)$, as described in the next section.

14.2.2 Dynamics of the Trajectory Basis Functions

As mentioned in Section 14.2 (and shown in Figure 14.1(c)), the TBFs can be seen as Gaussian functions following a *certain* dynamics. The question is: which dynamics should the TBFs follow to ensure a proper description of the underlying nuclear wave functions? Answering this question will allow us to define the three main methods employing TBFs for non-adiabatic quantum dynamics: variational Multi-configuration Gaussian (vMCG), multi-configuration Ehrenfest (MCE), or FMS.

In vMCG [33, 34, 43, 57, 82, 84], the equations of motion for all the Gaussian parameters are determined from the Dirac–Frenkel variational principle. This strategy ensures that the Gaussian dynamics will follow as closely as possible the one of the underlying nuclear wave functions. Hence, vMCG makes use of the best trajectories to minimize the overall number of TBFs required to describe the quantum dynamics accurately (see also Chapters 13 and 18 for more information on quantum trajectories and vMCG). This feature comes at the cost of the rather complex coupled dynamics of the TBFs, that is, the equations of motion for the parameters of a given Gaussian depend on the other Gaussians.

In MCE [37, 58, 60, 61], TBFs follow mean-field Ehrenfest trajectories (see Chapter 15 for a description of Ehrenfest dynamics). Depending on the version of the MCE algorithm considered, the dynamics of the Gaussian parameters can be coupled (as in vMCG) or independent, i.e., the propagation of the Gaussian parameters for one specific TBF does not depend on the other TBFs. MCE furthermore uses a time-independent width matrix, $\boldsymbol{\alpha}_k^{(\mu)}(t) \rightarrow \boldsymbol{\alpha}$. The use of Ehrenfest trajectories implies that, in regions of strong non-adiabatic coupling, a TBF can follow an average potential energy surface (composed of a linear combination of the coupled adiabatic potential energy surfaces, with weights given by electronic coefficients being propagated). An extension of the MCE, called *Ab Initio* Multiple Cloning (AIMC), proposes cloning trajectories on each adiabatic state after a non-adiabatic region, ensuring that the TBFs explore the different potential energy surfaces properly [38].

In FMS [3, 5, 7, 11, 22, 40, 41, 80], TBFs simply follow classical dynamics on their associated electronic state μ, and their width is frozen. For example, the position and momentum component ρ of the TBF k evolving in electronic state μ is given by Hamilton's equation of motion:

$$\frac{\partial \bar{R}_{k\rho}^{(\mu)}(t)}{\partial t} = \frac{\bar{P}_{k\rho}^{(\mu)}(t)}{M_\rho} \tag{14.14}$$

$$\frac{\partial \bar{P}_{k\rho}^{(\mu)}(t)}{\partial t} = -\frac{\partial E_\mu^{el}(\boldsymbol{R})}{\partial R_{k\rho}}\Big|_{R_{k\rho}=\bar{R}_{k\rho}^{(\mu)}(t)} . \tag{14.15}$$

The phase $\bar{\gamma}_k^{(\mu)}(t)$ is time-evolved by integrating the classical Lagrangian[3], based on semiclassical arguments,

$$\frac{\partial \bar{\gamma}_k^{(\mu)}(t)}{\partial t} = \sum_\rho^{3N_n} \frac{(\bar{P}_{k\rho}^{(\mu)}(t))^2}{2M_\rho} - E_\mu^{el}(\bar{\boldsymbol{R}}_k^{(\mu)}(t)) . \tag{14.16}$$

3 We note that the phase could also be absorbed into the time-dependent complex coefficients. However, this choice of phase allows an interaction picture for the complex coefficients to be defined, reducing their oscillations [11].

Hence, the dynamics of the TBFs in FMS becomes extremely simple. This simplicity comes at a cost, though, as an accurate description of the nuclear wave functions is likely to require a more significant number of classically-driven TBFs than vMCG ones (vMCG TBFs can explore a broader region of phase space). Also, a classical propagation for the TBFs prevents the description of tunneling processes, which can be accurately described in vMCG thanks to the variationally-propagated TBFs. FMS, however, overcome these difficulties by using a spawning algorithm that dynamically adapts the number of TBFs during the dynamics, ensuring that a proper number of TBFs are present in the case of non-adiabatic events. This algorithm will be described in Section 14.3.2.

Readers interested in more details about the MCE or the vMCG strategies can consult Chapter 13 as well as references [11, 37, 57]. In the following, we will focus our attention on the FMS framework.

14.3 Full Multiple Spawning

14.3.1 Full Multiple Spawning Equations

The FMS framework describes nuclear wave functions as linear combinations of frozen, classically-evolving Gaussians. The FMS dynamics starts with a certain number of *parent* TBFs, noted N_{ini}. To each parent is associated a given *branch* β (hence, there are N_{ini} branches). As described in Section 14.3.2, each TBF can create new TBFs via the spawning algorithm. For the moment, it is sufficient to note that each parent TBF β can have child TBFs (and that each child TBFs can spawn additional TBFs).

Hence, we can rewrite Eq. (14.17) in a slightly different way [11, 22], highlighting the link between parent and child TBFs more explicitly:

$$\begin{aligned}\Psi(\boldsymbol{r},\boldsymbol{R},t) &= \sum_{\beta}^{N_{ini}} \tilde{\Psi}_{\beta}(\boldsymbol{r},\boldsymbol{R},t) \\ &= \sum_{\beta}^{N_{ini}} \sum_{\nu}^{\infty} \sum_{k}^{N_{\nu}^{\beta}} C_{k\beta}^{(\nu)}(t)\,\tilde{\chi}_{k\beta}^{(\nu)}(\boldsymbol{R};\bar{\boldsymbol{R}}_{k\beta}^{(\nu)}(t),\bar{\boldsymbol{P}}_{k\beta}^{(\nu)}(t),\boldsymbol{\alpha},\bar{\gamma}_{k\beta}^{(\nu)}(t))\Phi_{\nu}(\boldsymbol{r};\boldsymbol{R})\ . \end{aligned} \tag{14.17}$$

The first line of Eq. (14.17) indicates that one can split the molecular wave function into (interacting) branches, labeled β, and each branch is initiated by a parent TBF. N_{ν}^{β} indicates the overall number of TBFs in electronic state ν in the branch of parent TBF β. For example, the parent TBF in the first branch ($\beta = 1$) in state ν would have a complex coefficient denoted $C_{11}^{(\nu)}(t)$, while the complex coefficient for the first child from this first parent TBF, in state μ, would be denoted by $C_{21}^{(\mu)}(t)$. We also note that the width matrix in Eq. (14.17) is now time-independent, and that the TBFs will follow classical trajectories, as described in Section 14.2.2. Employing the same protocol as described previously, we can derive equations of motion for the complex amplitudes:

$$\frac{d\boldsymbol{C}^{\mu}}{dt} = -i(\boldsymbol{S}^{-1})_{\mu\mu}\left[(\boldsymbol{H}_{\mu\mu} - i\dot{\boldsymbol{S}}_{\mu\mu})\boldsymbol{C}^{\mu} + \sum_{\nu\neq\mu}^{\infty}\boldsymbol{H}_{\mu\nu}\boldsymbol{C}^{\nu}\right]\ . \tag{14.18}$$

This equation is similar to Eq. (14.18), just focusing on the evolution of the complex coefficients in a given electronic state μ. The overlap matrices are described as before, $(\boldsymbol{S})_{k\beta,k'\beta'}^{\mu\mu} = \langle\tilde{\chi}_{k\beta}^{(\mu)}|\tilde{\chi}_{k'\beta'}^{(\mu)}\rangle_{\boldsymbol{R}}$ and $(\dot{\boldsymbol{S}})_{k\beta,k'\beta'}^{\mu\mu} = \langle\tilde{\chi}_{k\beta}^{(\mu)}|\frac{\partial}{\partial t}\tilde{\chi}_{k'\beta'}^{(\mu)}\rangle_{\boldsymbol{R}}$.

Let us now focus on the Hamiltonian matrix element between TBF k (from parent branch β) evolving in state μ and TBF k' (from parent branch β') evolving in state ν:

$$\begin{aligned}(\boldsymbol{H})^{\mu\nu}_{k\beta,k'\beta'} &= \langle \Phi_\mu \tilde{\chi}^{(\mu)}_{k\beta} | \hat{H} | \tilde{\chi}^{(\nu)}_{k'\beta'} \Phi_\nu \rangle_{\boldsymbol{r},\boldsymbol{R}} \\ &= \langle \tilde{\chi}^{(\mu)}_{k\beta} | \hat{T}_n | \tilde{\chi}^{(\nu)}_{k'\beta'} \rangle_{\boldsymbol{R}} \delta_{\mu\nu} + \langle \tilde{\chi}^{(\mu)}_{k\beta} | E^{el}_\nu | \tilde{\chi}^{(\nu)}_{k'\beta'} \rangle_{\boldsymbol{R}} \delta_{\mu\nu} \\ &\quad - \sum_\rho^{3N_n} \frac{1}{M_\rho} \langle \tilde{\chi}^{(\mu)}_{k\beta} | \langle \Phi_\mu | \frac{\partial}{\partial R_\rho} | \Phi_\nu \rangle_{\boldsymbol{r}} \frac{\partial}{\partial R_\rho} | \tilde{\chi}^{(\nu)}_{k'\beta'} \rangle_{\boldsymbol{R}} \\ &\quad - \sum_\rho^{3N_n} \frac{1}{2M_\rho} \langle \tilde{\chi}^{(\mu)}_{k\beta} | \langle \Phi_\mu | \frac{\partial^2}{\partial R^2_\rho} | \Phi_\nu \rangle_{\boldsymbol{r}} | \tilde{\chi}^{(\nu)}_{k'\beta'} \rangle_{\boldsymbol{R}} \,. \end{aligned} \tag{14.19}$$

If the two TBFs are evolving on the same electronic state, their coupling is determined by the first two terms on the right-hand side of Eq. (14.19), that is, the terms containing the nuclear kinetic energy operator and the electronic energy, as well as the diagonal contribution ($\mu = \nu$) of the last term of the right-hand side, containing, in this case, diagonal Born–Oppenheimer corrections (DBOCs)

$$D_{\mu\mu}(\boldsymbol{R}) = \langle \Phi_\mu | \frac{\partial^2}{\partial R^2_\rho} | \Phi_\mu \rangle_{\boldsymbol{r}} \,. \tag{14.20}$$

These three terms are responsible for the description of *intrastate* couplings (Figure 14.2(b)) and will be central to adequately describing the dynamics of the nuclear amplitude on a given electronic state.

If the two TBFs are in different electronic states, they will interact via the third term on the r.h.s of Eq. (14.19) that depends on the non-adiabatic coupling vectors,

$$\boldsymbol{d}_{\mu\nu}(\boldsymbol{R}) = \langle \Phi_\mu | \frac{\partial}{\partial \boldsymbol{R}} | \Phi_\nu \rangle_{\boldsymbol{r}} \,, \tag{14.21}$$

and the off-diagonal contribution in the last term of the r.h.s, containing the second-order non-adiabatic coupling:

$$D_{\mu\nu}(\boldsymbol{R}) = \langle \Phi_\mu | \frac{\partial^2}{\partial R^2_\rho} | \Phi_\nu \rangle_{\boldsymbol{r}} \,. \tag{14.22}$$

The non-adiabatic coupling terms provide *interstate* couplings (Figure 14.2(a)) and are responsible for the transfer of nuclear amplitude between electronic states due to nuclear motion, i.e., the non-Born–Oppenheimer effects. Figure 14.2 provides a schematic representation of the couplings between three TBFs, where the red areas indicate interstate couplings and the blue area refers to

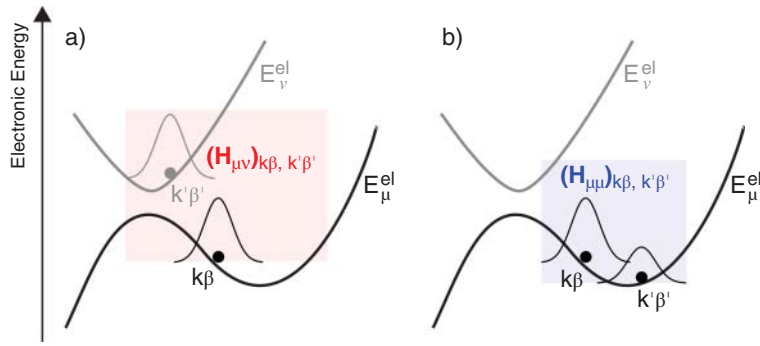


Figure 14.2 Schematic representation of an interstate (a) and an intrastate (b) coupling between two TBFs.

an intrastate coupling. It is worth pointing out that FMS would tend towards an exact solution of the time-dependent molecular Schrödinger equation if we were to use a large number of TBFs and could compute all Hamiltonian matrix elements (Eq. (14.19)) exactly.

14.3.2 Spawning Algorithm

Up to this point, we have considered that we can define a certain number of parent TBFs (N_{ini}), each having a fixed number of child TBFs – leading to an overall number of TBFs in state μ for the branch β given by N_μ^β. The description of a crucial ingredient of FMS is still missing: the spawning algorithm, which permits the number of TBFs to be adapted dynamically during the non-adiabatic dynamics. The idea of spawning TBFs is as follows: one starts with a swarm of N_{ini} coupled parent TBFs, whose distribution and complex coefficients are chosen such that they reproduce an initial nuclear wave packet (for example, generated after photo-excitation of a molecule to a given excited electronic state). A strategy would be to use a considerable number of TBFs from time $t = 0$ and hope that the TBFs will follow the nuclear wave packet(s) during the excited-state dynamics. Instead, an FMS dynamics starts with a rather small number of TBFs (enough to describe the early stage of the dynamics) but will give the opportunity to each TBF (being it parent or child) to spawn new TBFs on a coupled electronic state whenever non-adiabatic regions are encountered. For example, if one parent TBF reaches a region of strong non-adiabaticity, it can *spawn* a child TBF onto the coupled electronic state. By *spawn*, we mean here that a new TBF is created, which implies (*i*) that the number of TBFs in each electronic state for each branch can change over time, $N_\mu^\beta \rightarrow N_\mu^\beta(t)$, and (*ii*) that the size of the matrices in Eq. (14.18) will change over time too. Hence, one should make sure that these new TBFs are introduced in the dynamics only when the coupling between electronic states is still small to ensure a smooth propagation of the amplitudes: this is the role of the spawning algorithm.

The main steps of the spawning algorithm are represented in Figure 14.3. t_0 corresponds to the time where a TBF k (from branch β) in state ν is approaching a region of strong non-adiabatic coupling. At each nuclear time step, the TBF monitors an effective non-adiabatic coupling with all other electronic states μ, $\Lambda_{\mu\nu}^{eff}(\bar{\boldsymbol{R}}_{k\beta}^{(\nu)})$. Different definitions for the effective coupling have been proposed, and the most commonly employed one in the adiabatic representation is the modulus of the non-adiabatic coupling vectors, $\Lambda_{\mu\nu}^{eff}(\bar{\boldsymbol{R}}_{k\beta}^{(\nu)}) = |\boldsymbol{d}_{\mu\nu}(\bar{\boldsymbol{R}}_{k\beta}^{(\nu)})|$, or the projection of the non-adiabatic coupling vectors on the TBF classical velocities, $\Lambda_{\mu\nu}^{eff}(\bar{\boldsymbol{R}}_{k\beta}^{(\nu)}) = |\boldsymbol{d}_{\mu\nu}(\bar{\boldsymbol{R}}_{k\beta}^{(\nu)}) \cdot \dot{\bar{\boldsymbol{R}}}_{k\beta}^{(\nu)}|$. The value of $\Lambda_{\mu\nu}^{eff}(\bar{\boldsymbol{R}}_{k\beta}^{(\nu)}(t))$ along the TBF is represented in Figure 14.3 by the thin blue line. When the TBF enters a region (blue area in Figure 14.3) where $\Lambda_{\mu\nu}^{eff}(\bar{\boldsymbol{R}}_{k\beta}^{(\nu)}(t))$ is larger than a predefined threshold (dashed horizontal line in Figure 14.3), the dynamics enters the so-called *spawning mode*. From this time on, called the entry time t_{entry}, the propagation of the complex amplitudes is suspended (this is represented in the central panel of Figure 14.3 by the use of dotted lines for the TBFs) and one simply propagates the TBF classically, until a maximum in $\Lambda_{\mu\nu}^{eff}(\bar{\boldsymbol{R}}_{k\beta}^{(\nu)}(t))$ is recorded. If such a maximum is detected, a new TBF is created (at time t_{spawn}) on the coupled state μ, with a complex coefficient set to zero (flat dotted line in Figure 14.3). From this point, both TBFs are back-propagated in time from t_{spawn} to t_{entry}. It is important to realize that the newly created TBF will have a different dynamics than the parent TBF as it evolves on electronic state μ. Once t_{entry} is reached, the dynamics leaves the spawning mode and the FMS dynamics can resume, with equations of motion for the complex amplitudes augmented by a new function. The original TBF k in electronic state ν will meet the newly-created TBF evolving in state μ exactly in the region of strong non-adiabatic coupling, ensuring a proper description of the nuclear quantum dynamics and a potential transfer of amplitude between the two TBFs. More information on the spawning algorithm can be found in Refs. [5, 11].

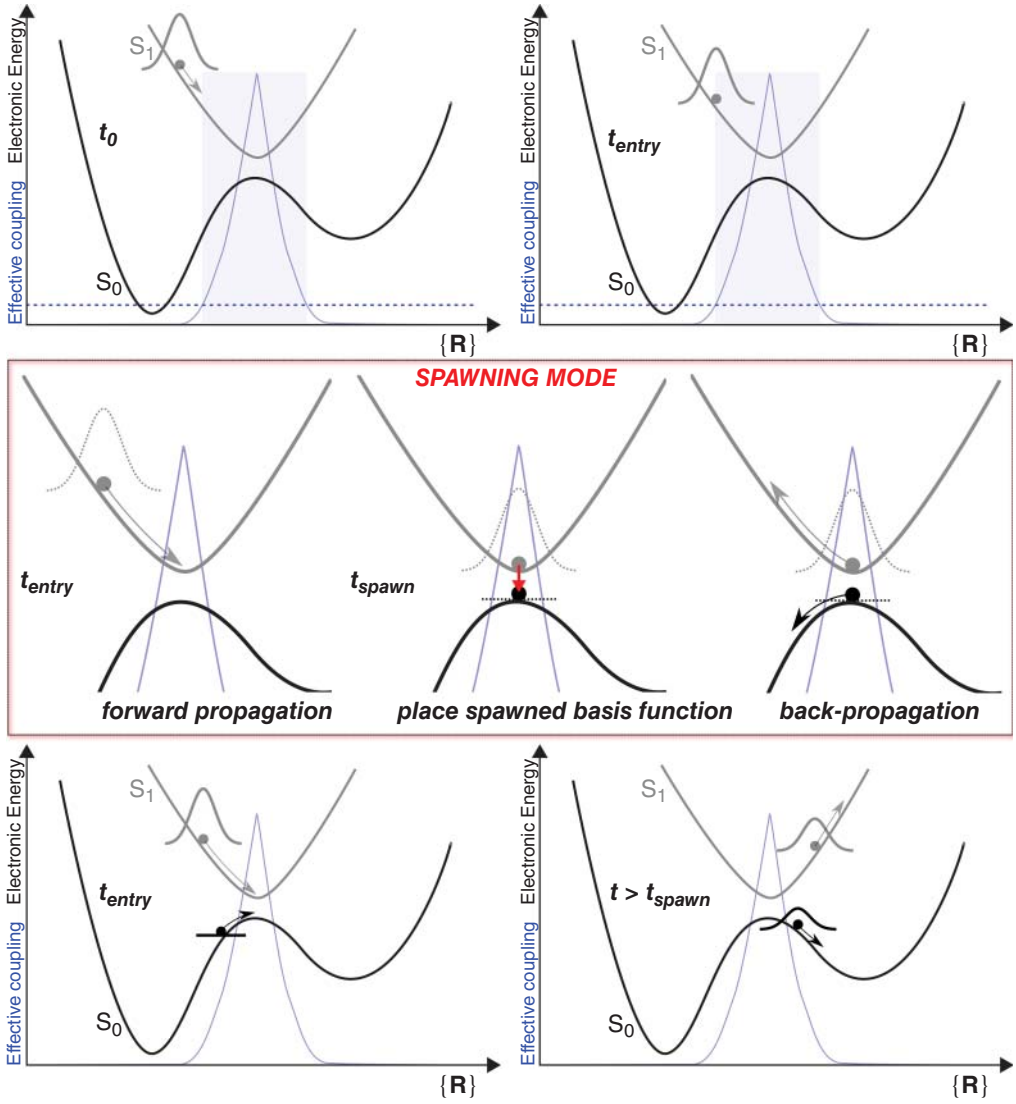


Figure 14.3 Schematic representation of the spawning algorithm: the entry in a region of strong non-adiabaticity (upper panel), the spawning mode (central panel), and the subsequent dynamics upon spawning of a new TBF. Adapted with permission from Ref. [11]. Copyright (2018) American Chemical Society.

The spawning algorithm described above is rather simple and only allows for interstate spawnings. Other, different, strategies for spawning TBFs have been described in the literature [5, 6, 86], also for the specific case of tunneling where intrastate spawning is required [4].

14.4 Extending Full Multiple Spawning

The FMS equations of motion introduced up to now account for the non-adiabatic dynamics of nuclear wave packets in an *in principle* exact way. TBFs are coupled via intra- and interstate

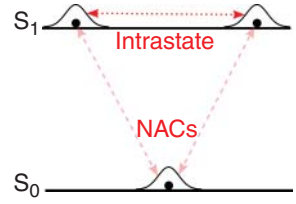


Figure 14.4 Coupling pattern between TBFs in the original version of FMS. Intrastate couplings (red dotted line) are mediated by the nuclear kinetic energy operator, the electronic energy, and the DBOCs. Interstate couplings (light red dashed arrow) originate from first- and second-order non-adiabatic couplings.

terms, which depend on the nuclear kinetic energy operator, electronic energy, and non-adiabatic coupling terms of first and second order (Figure 14.4). Hence, the FMS framework, in its original flavor, allows for the description of non-radiative de-excitation of molecules via internal conversion pathways. There are, however, other different physical processes related to excited states one may want to describe such as intersystem crossings (transfer of nuclear wave packets between states with different spin multiplicities due to spin–orbit coupling) or the explicit interaction of a molecule with an external electromagnetic field. Such processes can easily be included in FMS by adding to the molecular Hamiltonian (Eq. (14.2)) any desired interaction Hamiltonians. In an FMS picture, these new interactions will simply add new couplings between TBFs and require additional spawning events. In the following, we detail two extensions of FMS that propose including the effect of an explicit external field and spin–orbit coupling.

14.4.1 External Field in Full Multiple Spawning

Most non-adiabatic molecular dynamics simulations start at time t_0 with a nuclear wave packet placed in a chosen excited electronic state. This transfer of a ground-state nuclear wave function into a given excited electronic state is expected to mimic the result of a photo-excitation process conducted with an ultrashort laser pulse. But how can we simulate this photo-excitation process explicitly in FMS?[4] A simple solution is to explicitly include the coupling between a time-dependent external field and the molecular dipole moment. In other words, we need to supplement the original molecular Hamiltonian defined in Eq. (14.2) by a new interaction term between the molecule and the time-dependent external (electric) field $\underline{\boldsymbol{E}}(t)$,

$$\hat{H}(\boldsymbol{r},\boldsymbol{R}) \rightarrow \hat{\mathcal{H}}(\boldsymbol{r},\boldsymbol{R},t) = \hat{H}(\boldsymbol{r},\boldsymbol{R}) - \underline{\hat{\boldsymbol{\mu}}}(\boldsymbol{r},\boldsymbol{R}) \cdot \underline{\boldsymbol{E}}(t) \,, \tag{14.23}$$

where $\underline{\hat{\boldsymbol{\mu}}}(\boldsymbol{r},\boldsymbol{R}) = \underline{\hat{\boldsymbol{\mu}}}^e(\boldsymbol{r}) + \underline{\hat{\boldsymbol{\mu}}}^n(\boldsymbol{R})$ is the molecular dipole moment, composed of an electronic and a nuclear part. In the following, underlined bold symbols indicate 3D vectors and $\hat{\mathcal{H}}$ designates a modified molecular Hamiltonian. The presence of the coupling with an external field will induce new couplings between TBFs. We can substitute the molecular Hamiltonian by the modified one defined in Eq. (14.23) in the time-dependent Schrödinger equation, and re-derive the FMS equations (using the recipe described in Section 14.3). The equations of motion for the complex coefficients will be similar to Eq. (14.18), except that now the Hamiltonian matrix elements

4 We note that the effect of an external field in FMS has also been included in a Floquet-type approach [30, 31].

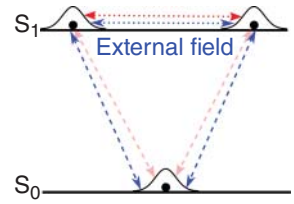


Figure 14.5 Coupling pattern between TBFs in the eXternal Field Full Multiple Spawning (XFFMS) technique. In comparison to the original FMS (Figure 14.4), the external field causes additional intrastate (blue dotted arrow) and interstate (blue dashed arrows) couplings between TBFs.

contain new terms originating from the coupling with the external field:

$$\begin{aligned}(\mathcal{H})^{\mu\nu}_{k\beta,k'\beta'} &= \langle \Phi_\mu \tilde{\chi}^{(\mu)}_{k\beta} | \hat{H} | \tilde{\chi}^{(\nu)}_{k'\beta'} \Phi_\nu \rangle_{\boldsymbol{r},\boldsymbol{R}} - \langle \Phi_\mu \tilde{\chi}^{(\mu)}_{k\beta} | \underline{\hat{\boldsymbol{\mu}}} | \tilde{\chi}^{(\nu)}_{k'\beta'} \Phi_\nu \rangle_{\boldsymbol{r},\boldsymbol{R}} \cdot \underline{\boldsymbol{E}}(t) \\ &= \langle \Phi_\mu \tilde{\chi}^{(\mu)}_{k\beta} | \hat{H} | \tilde{\chi}^{(\nu)}_{k'\beta'} \Phi_\nu \rangle_{\boldsymbol{r},\boldsymbol{R}} - \langle \tilde{\chi}^{(\mu)}_{k\beta} | \underline{\hat{\boldsymbol{\mu}}}^n | \tilde{\chi}^{(\nu)}_{k'\beta'} \rangle_{\boldsymbol{R}} \cdot \underline{\boldsymbol{E}}(t) \delta_{\mu\nu} \\ &\quad - \langle \tilde{\chi}^{(\mu)}_{k\beta} | \langle \Phi_\mu | \underline{\hat{\boldsymbol{\mu}}}^e | \Phi_\nu \rangle_{\boldsymbol{r}} | \tilde{\chi}^{(\nu)}_{k'\beta'} \rangle_{\boldsymbol{R}} \cdot \underline{\boldsymbol{E}}(t) \, . \end{aligned} \tag{14.24}$$

Comparing with the original FMS Hamiltonian matrix elements (Eq. (14.19)), we observe that the term depending on the nuclear component of the dipole moment (second term on the right-hand side of Eq. (14.24)) adds an extra intrastate coupling between TBFs (dotted blue arrow in Figure 14.5). If the two TBFs are in the same electronic state, the electronic part of the dipole moment (last two terms on the right-hand side of Eq. (14.24)) contributes an intrastate coupling term (dotted blue arrow in Figure 14.5) via the electronic dipole moment

$$\underline{\boldsymbol{\mu}}^e_{\mu\mu}(\boldsymbol{R}) = \langle \Phi_\mu | \underline{\hat{\boldsymbol{\mu}}}^e | \Phi_\mu \rangle_{\boldsymbol{r}} \, . \tag{14.25}$$

If, on the other hand, the TBFs are in two different electronic states, the external field can couple them via the corresponding transition dipole moment,

$$\underline{\boldsymbol{\mu}}^e_{\mu\nu}(\boldsymbol{R}) = \langle \Phi_\mu | \underline{\hat{\boldsymbol{\mu}}}^e | \Phi_\nu \rangle_{\boldsymbol{r}} \, . \tag{14.26}$$

It is this interstate coupling term (dashed blue arrow in Figure 14.5) that is responsible for the transfer of amplitude between electronic state mediated by the external field (or laser pulse).

The spawning algorithm needs to be adapted to ensure that TBFs are present in regions of coupling induced by the external field. As the external field is known in advance, a simple way to extend the spawning algorithm is to induce a spawning event to the coupled state whenever $\underline{\boldsymbol{E}}(t)$ shows an extremum. The overall method is called *eXternal Field Full Multiple Spawning* (XFFMS), and allows for an *in silico* photo-chemical experiment: an XFFMS dynamics starts with a molecule in its ground electronic state coupled with an explicit time-dependent external field, eventually leading to amplitude transfer to a coupled electronic state and followed by non-adiabatic relaxation towards the ground electronic state. XFFMS also permits the simulation of more advanced pump–probe experiments by defining a more complex laser pulse pattern [46, 49]. Additional details can be found in Refs. [11, 44, 47].

14.4.2 Spin-Orbit Coupling in Full Multiple Spawning

FMS, in its original formulation, focuses on non-adiabatic transitions involving internal conversion processes only (transfer of nuclear amplitude between electronic states sharing the same spin multiplicity), but is unable to describe intersystem crossings that would require the inclusion of couplings between states of different spin multiplicities (singlet and triplet states, for example).

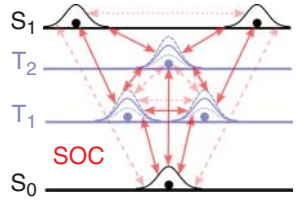


Figure 14.6 Coupling pattern between TBFs in the Generalized Full Multiple Spawning (GFMS) technique. As an example, two singlet states (black) are coupled to two triplet states (blue). A triplet TBF is depicted with three Gaussian shapes, symbolizing the three different complex amplitudes – one for each spin sublevel. Red arrows indicate the couplings between TBFs induced by spin–orbit coupling. Couplings from the original FMS formulation are depicted by light red arrows (intrastate: dotted; interstate: dashed). (Note that for clarity we do not indicate all the possible couplings between spin sublevels.)

This omission is the direct result of the use of a non-relativistic formulation of quantum mechanics, that is, the use of the time-dependent Schrödinger equation Eq. (14.1) as the starting point of our derivation. The relativistic effect called spin–orbit coupling [39, 54] can, however, still be included in the Schrödinger equation by employing the spin–orbit contribution to the so-called Breit–Pauli Hamiltonian. Modifying the molecular Hamiltonian (Eq. (14.2)) to include the spin–orbit coupling contribution, we obtain:

$$\hat{H}(\boldsymbol{r},\boldsymbol{R}) \rightarrow \hat{\mathcal{H}}(\boldsymbol{x},\boldsymbol{R}) = \hat{H}(\boldsymbol{r},\boldsymbol{R}) + \hat{H}_{SOC}(\boldsymbol{x},\boldsymbol{R}) \,, \tag{14.27}$$

where $\boldsymbol{x} = (\boldsymbol{r}, s)$ (we indicate here the spin component for the electronic degrees of freedom only). As was done in the original derivation of FMS, we will use in the following spin-free adiabatic electronic states, leading to a dynamics in a spin-diabatic representation [20, 36, 54]. We need, however, to indicate explicitly the spin multiplicity and spin-projection eigenvalue for each electronic state considered. Hence, we will note in the following the spin-free, adiabatic electronic wave functions as $\Phi_\nu^{M_{S_\nu}}(\boldsymbol{x};\boldsymbol{R})$, with S_ν the total spin of electronic state ν and M_{S_ν} its spin-projection eigenvalue.[5] Using this notation, the FMS version of the BH representation (Eq. (14.17)) becomes

$$\Psi(\boldsymbol{x},\boldsymbol{R},t) = \sum_{\beta}^{N_{ini}} \sum_{\nu}^{\infty} \sum_{k}^{N_{\nu,M_{S_\nu}}^{\beta}} C_{k\beta}^{(\nu,M_{S_\nu})}(t)\, \tilde{\gamma}_{k\beta}^{(\nu,M_{S_\nu})}(\boldsymbol{R};\bar{\boldsymbol{R}}_{k\beta}^{(\nu,M_{S_\nu})}(t),\bar{\boldsymbol{P}}_{k\beta}^{(\nu,M_{S_\nu})}(t),\boldsymbol{\alpha},\bar{\gamma}_{k\beta}^{(\nu,M_{S_\nu})}(t))\Phi_\nu^{M_{S_\nu}}(\boldsymbol{x};\boldsymbol{R}) \,. \tag{14.28}$$

It is important to note that Eq. (14.28) and Eq. (14.17) are equivalent – the former simply includes additional information about the spin-state of a given electronic state. In this formulation, a TBF evolving on a singlet state ν ($S_\nu = 0$) will have a label $M_{S_\nu} = 0$. In the specific case of a triplet electronic state μ ($S_\mu = 1$), we have to define TBFs for each triplet sublevel, i.e., for $M_{S_\mu} = -1$, $M_{S_\mu} = 0$, and $M_{S_\mu} = 1$. Hence, a Gaussian function evolving on a triplet state will have three distinct complex coefficients, one for each triplet sublevel (see Figure 14.6).

Inserting Eq. (14.28) into the time-dependent Schrödinger equation, using the modified Hamiltonian defined in Eq. (14.27), we obtain – after some algebra similar to what was described in Section 14.3 – a set of equations of motion for the complex amplitudes whose form is reminiscent to the one of FMS (Eq. (14.18)). Two main differences have to be noted: (*i*) the complex coefficients have an additional label defining the spin sublevel of the electronic state to which they are attached, and (*ii*) the TBFs suffer additional couplings coming from the spin–orbit coupling term in Eq. (14.27).

5 For a singlet state, we have $S = 0$ and $M_S = 0$. A triplet state is characterized by $S = 1$ and $M_S = -1$, 0, or 1.

Let us investigate a Hamiltonian matrix element between TBF k (from parent branch β) evolving in state μ with spin-projection M_{S_μ} and TBF k' (from parent branch β') evolving in state ν with spin-projection M_{S_ν}:

$$\begin{aligned}(\mathcal{H})_{k\beta,k'\beta'}^{\mu M_{S_\mu},\nu M_{S_\nu}} &= \langle \Phi_\mu^{M_{S_\mu}} \tilde{\chi}_{k\beta}^{(\mu,M_{S_\mu})} | \hat{H} | \tilde{\chi}_{k'\beta'}^{(\nu,M_{S_\nu})} \Phi_\nu^{M_{S_\nu}} \rangle_{\boldsymbol{r},\boldsymbol{R}} \delta_{M_{S_\mu} M_{S_\nu}} \\ &\quad + \langle \Phi_\mu^{M_{S_\mu}} \tilde{\chi}_{k\beta}^{(\mu,M_{S_\mu})} | \hat{H}_{SOC} | \tilde{\chi}_{k'\beta'}^{(\nu,M_{S_\nu})} \Phi_\nu^{M_{S_\nu}} \rangle_{\boldsymbol{x},\boldsymbol{R}} \,. \end{aligned} \tag{14.29}$$

The coupling between TBFs generated by the spin–orbit coupling Hamiltonian (last term on the right-hand side of Eq. (14.29)) can be rewritten as

$$\begin{aligned}(H^{SOC})_{k\beta,k'\beta'}^{\mu M_{S_\mu},\nu M_{S_\nu}} &= \langle \Phi_\mu^{M_{S_\mu}} \tilde{\chi}_{k\beta}^{(\mu,M_{S_\mu})} | \hat{H}_{SOC} | \tilde{\chi}_{k'\beta'}^{(\nu,M_{S_\nu})} \Phi_\nu^{M_{S_\nu}} \rangle_{\boldsymbol{x},\boldsymbol{R}} \\ &= \langle \tilde{\chi}_{k\beta}^{(\mu,M_{S_\mu})} | \langle \Phi_\mu^{M_{S_\mu}} | \hat{H}_{SOC} | \Phi_\nu^{M_{S_\nu}} \rangle_{\boldsymbol{x}} | \tilde{\chi}_{k'\beta'}^{(\nu,M_{S_\nu})} \rangle_{\boldsymbol{R}} \,, \end{aligned} \tag{14.30}$$

highlighting the fact that the spin–orbit coupling will induce a rather complex coupling pattern between TBFs. In short: in addition to the already existing non-relativistic couplings, spin–orbit coupling can provide an intrastate coupling for TBFs on the same electronic state and same spin sublevel, an interstate coupling for TBFs on the same electronic state but with different spin-sublevels, an interstate coupling between TBFs on different electronic states but with the same spin multiplicity and different spin-sublevels, and finally an interstate coupling for TBFs in different electronic states with different spin multiplicity. Figure 14.6 provides a schematic representation of the different couplings induced by spin–orbit coupling.

The spawning algorithm is adapted to detect regions of important spin–orbit coupling, and the effective coupling described in Section 14.3.2 not only monitors non-adiabatic couplings, but also the strength of spin–orbit coupling (in a spin-diabatic picture) felt by a TBF k evolving in state ν with M_{S_ν}:

$$\Lambda_{\mu\nu}^{eff}(\bar{\boldsymbol{R}}_k(t)) = \frac{\left(\sum_{M_{S_\mu}=-S_\mu}^{S_\mu} \sum_{M_{S_\nu}=-S_\nu}^{S_\nu} |\langle \Phi_\mu^{M_{S_\mu}}(\bar{\boldsymbol{R}}_k) | \hat{H}_{SOC} | \Phi_\nu^{M_{S_\nu}}(\bar{\boldsymbol{R}}_k) \rangle_{\boldsymbol{x}}|^2 \right)^{1/2}}{|E_\nu^{el}(\bar{\boldsymbol{R}}_k) - E_\mu^{el}(\bar{\boldsymbol{R}}_k)|} \,. \tag{14.31}$$

This overall extension of FMS to both intersystem crossing and internal conversion processes is called *Generalized Full Multiple Spawning* (GFMS) and more details can be found in Refs. [11, 12]. An alternative implementation of spin–orbit coupling in FMS is presented in Refs. [16, 17].

14.5 *Ab Initio* Multiple Spawning

14.5.1 From Full- to *Ab Initio* Multiple Spawning

In the following, we introduce the AIMS method, which translates the FMS framework to the excited-state dynamics of molecules in their full configuration space. The formalism of FMS presented above, as well as its extensions, is *in principle* exact given a sufficiently large number of TBFs and the possibility to compute all the Hamiltonian matrix elements (Eqs. (14.19) for the original FMS). While these conditions can be achieved for model systems in low dimensions, simulating the excited-state dynamics of molecules implies the introduction of two central approximations[6] that will define the AIMS method [3, 4, 22, 40].

6 In addition to the AIMS approximations described in this paragraph, we note that a practical application of AIMS requires the definition of a subset of coupled electronic states, i.e., $\sum_\mu^\infty \rightarrow \sum_\mu^{N_s}$ with N_s being the number of electronic states explicitly considered for the dynamics.

The first approximation, called the *saddle-point approximation*, is related to the calculation of the Hamiltonian matrix elements. In Eq. (14.19), some integrals are performed over the nuclear coordinates while containing electronic structure quantities like the electronic energies or the non-adiabatic couplings. Calculating these integrals would imply knowing these electronic structure quantities over the entire configuration space – in other words, we would have to precompute all electronic structure quantities in advance, for all possible nuclear displacements, as done in numerically-exact wave packet propagation on a grid. The cost associated with this process would hamper the simulation of molecules in their full configuration space, and one needs to approximate the Hamiltonian matrix elements.

Let us first consider a typical integral between TBF k in μ and TBF k' in ν containing the arbitrary electronic structure quantity $\theta_{\mu\nu}(\boldsymbol{R}) = \langle\Phi_\mu|\hat{\theta}(\boldsymbol{R})|\Phi_\nu\rangle_{\boldsymbol{r}}$:

$$(\boldsymbol{\theta})^{\mu\nu}_{k\beta,k'\beta'} = \langle\tilde{\chi}^{(\mu)}_{k\beta}|\langle\Phi_\mu|\hat{\theta}|\Phi_\nu\rangle_{\boldsymbol{r}}|\tilde{\chi}^{(\nu)}_{k'\beta'}\rangle_{\boldsymbol{R}} \,. \tag{14.32}$$

We now define the centroid position of the product of the two Gaussian functions,

$$\bar{\boldsymbol{R}}^{(\mu\nu)}_{k\beta,k'\beta'} = \frac{(\bar{\boldsymbol{R}}^{(\mu)}_{k\beta} + \bar{\boldsymbol{R}}^{(\nu)}_{k'\beta'})}{2} \,, \tag{14.33}$$

and Taylor-expand the electronic structure quantity $\theta_{\mu\nu}(\boldsymbol{R})$ around this position:

$$\begin{aligned}\theta_{\mu\nu}(\boldsymbol{R}) = \theta_{\mu\nu}(\bar{\boldsymbol{R}}^{(\mu\nu)}_{k\beta,k'\beta'}) + \sum_{\rho}^{3N_n}(R_\rho - \bar{R}^{(\mu\nu)}_{\rho,k\beta,k'\beta'})\left.\frac{\partial\theta_{\mu\nu}(\boldsymbol{R})}{\partial R_\rho}\right|_{R_\rho=\bar{R}^{(\mu\nu)}_{\rho,k\beta,k'\beta'}} \\ + \frac{1}{2}\sum_{\rho\rho'}^{3N_n}(R_\rho - \bar{R}^{(\mu\nu)}_{\rho,k\beta,k'\beta'})\frac{\partial^2\theta_{\mu\nu}(\boldsymbol{R})}{\partial R_\rho\partial R_{\rho'}}|_{R_\rho=\bar{R}^{(\mu\nu)}_{\rho,k\beta,k'\beta'},R_{\rho'}=\bar{R}^{(\mu\nu)}_{\rho',k\beta,k'\beta'}}(R_{\rho'} - \bar{R}^{(\mu\nu)}_{\rho'k\beta,k'\beta'}) + \ldots\end{aligned} \tag{14.34}$$

Owing to the locality of Gaussian functions, we can consider that the electronic structure quantity of interest only slightly varies within the region where the two TBFs overlap substantially. Hence, we could only keep the term of order zero (containing the electronic structure quantity evaluated at the centroid position) in Eq (14.34) and neglect the $\boldsymbol{R}$ dependence of $\theta_{\mu\nu}(\boldsymbol{R})$ by neglecting all the terms containing derivatives:

$$\theta_{\mu\nu}(\boldsymbol{R}) \approx \theta_{\mu\nu}(\bar{\boldsymbol{R}}^{(\mu\nu)}_{k\beta,k'\beta'}) \,. \tag{14.35}$$

As a result of this approximation, the integrals containing electronic structure quantities (of the form of Eq. (14.32)) in the Hamiltonian matrix elements can be approximated by

$$\begin{aligned}(\boldsymbol{\theta})^{\mu\nu}_{k\beta,k'\beta'} &= \langle\tilde{\chi}^{(\mu)}_{k\beta}|\langle\Phi_\mu|\hat{\theta}|\Phi_\nu\rangle_{\boldsymbol{r}}|\tilde{\chi}^{(\nu)}_{k'\beta'}\rangle_{\boldsymbol{R}} \\ &\approx \theta_{\mu\nu}(\bar{\boldsymbol{R}}^{(\mu\nu)}_{k\beta,k'\beta'})\langle\tilde{\chi}^{(\mu)}_{k\beta}|\tilde{\chi}^{(\nu)}_{k'\beta'}\rangle_{\boldsymbol{R}} \,.\end{aligned} \tag{14.36}$$

This approximation, called *saddle point approximation of order zero* (SPA0), substantially simplifies the calculation of coupling matrix elements between TBFs. Applying the SPA0 to Eq. (14.19) and neglecting the second-order couplings leads to the AIMS Hamiltonian matrix elements:

$$\begin{aligned}(\boldsymbol{H})^{\mu\nu}_{k\beta,k'\beta'} \approx \langle\tilde{\chi}^{(\mu)}_{k\beta}|\hat{T}_n|\tilde{\chi}^{(\nu)}_{k'\beta'}\rangle_{\boldsymbol{R}}\delta_{\mu\nu} + E^{el}_\nu(\bar{\boldsymbol{R}}^{(\mu\nu)}_{k\beta,k'\beta'})\langle\tilde{\chi}^{(\mu)}_{k\beta}|\tilde{\chi}^{(\nu)}_{k'\beta'}\rangle_{\boldsymbol{R}}\delta_{\mu\nu} \\ - \sum_{\rho}^{3N_n}\frac{1}{M_\rho}(\boldsymbol{d}_{\mu\nu}(\bar{\boldsymbol{R}}^{(\mu\nu)}_{k\beta,k'\beta'}))_\rho\langle\tilde{\chi}^{(\mu)}_{k\beta}\left|\frac{\partial}{\partial R_\rho}\right|\tilde{\chi}^{(\nu)}_{k'\beta'}\rangle_{\boldsymbol{R}} \,.\end{aligned} \tag{14.37}$$

The SPA0 is not only valid for FMS, but can also be used in XFFMS (to approximate the matrix elements containing the dipole and transition dipole moments in Eq. (14.24)) or in GFMS (to

approximate the matrix elements containing spin–orbit coupling in Eq. (14.29)), leading to the methods named XFAIMS and GAIMS, respectively.

The second possible approximation when treating molecular systems is called the *independent first generation approximation* (IFGA). As detailed previously, the FMS dynamics is initiated at time $t = 0$ by considering a set of N_{ini} coupled parent TBFs that reproduce the initial nuclear wave packet in a given (excited) electronic state (see Eq. (14.17)). However, a *high-dimensional* nuclear wave packet will usually rapidly spread in phase space as soon as dynamics in the excited state starts, meaning that the originally coupled parent TBFs will rapidly move away from each other and become uncoupled. Hence, we could approximate that the parent TBFs are, from the beginning of the dynamics, uncoupled. Within this approximation, we are allowed to sample the parent TBFs independently (for example from a Wigner distribution), set their initial amplitude to 1, and propagate them independently. In other words, the complex coefficients of TBFs evolving in the branch β are not coupled to those of TBFs evolving in a different branch β' [5, 22]. Mathematically, this approximation can be enforced by applying the conditions $(\boldsymbol{S}_{\mu\nu})_{k\beta,k'\beta'} \approx (\boldsymbol{S}_{\mu\nu})_{k,k'}\delta_{\beta\beta'}$ and $(\boldsymbol{H}_{\mu\nu})_{k\beta,k'\beta'} \approx (\boldsymbol{H}_{\mu\nu})_{k,k'}\delta_{\beta\beta'}, \forall \mu, \nu$.

Applying the SPA0 and the IFGA to the FMS framework defines the AIMS method (see Figure 14.7, left panel). More information about these approximations can be found in Refs. [5, 11].

14.5.2 Testing the Approximations of *Ab Initio* Multiple Spawning

The AIMS method relies on the validity of the SPA0 and the IFGA. Different works have focused on assessing the quality of these approximations [2–4, 6, 22, 40, 72], and we describe here a stringent test based on the use of an ultrashort light pulse within XFFMS to generate highly interfering nuclear wave packets [44]. More specifically, this test focuses on the photo-excitation of lithium hydride (LiH), a one-dimensional system. By sending an ultrashort pulse on this molecule, originally in its ground electronic state (S_0), part of the nuclear amplitude is transferred to the first electronic state (S_1) and forms a nuclear wave packet. Due to the different shape of the potential

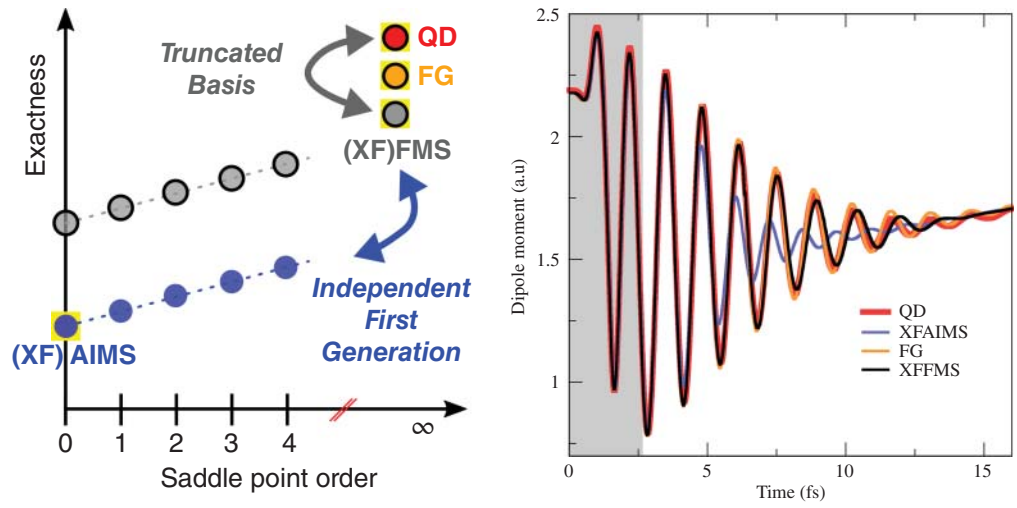


Figure 14.7 Left panel: Scheme representing the different approximations applied to QD to reach (XF)FMS and (XF)AIMS. Right panel: Time-dependent dipole moment as computed with XFFMS (black), XFAIMS (blue), fixed Gaussian (FG, orange), and exact QD (red). The gray area represents the time window during which the ultrashort pulse is switched on. Adapted from Ref. [44] with the permission of AIP Publishing.

energy curves between S_0 and S_1, the nuclear wave packet formed on S_1 will immediately leave the Franck–Condon region. A challenging observable one could use to assess the quality of the XFFMS (within the SPA0 and the IFGA) simulation is the time-dependent dipole moment, which, in an FMS notation, reads

$$\langle\Psi|\hat{\boldsymbol{\mu}}|\Psi\rangle_{\boldsymbol{r},\boldsymbol{R}} = \sum_{\beta\beta'}^{N_{ini}} \sum_{\mu\nu}^{\infty} \left[\sum_{k}^{N_\mu^\beta(t)} \sum_{k'}^{N_\nu^{\beta'}(t)} (C_{k\beta}^{(\mu)}(t))^* C_{k'\beta'}^{(\nu)}(t) [\langle \tilde{\chi}_{k\beta}^{(\mu)} | \underline{\hat{\boldsymbol{\mu}}}^n | \tilde{\chi}_{k'\beta'}^{(\nu)} \rangle_{\boldsymbol{R}} \delta_{\mu\nu} + \langle \tilde{\chi}_{k\beta}^{(\mu)} | \langle \Phi_\mu | \underline{\hat{\boldsymbol{\mu}}}^e | \Phi_\nu \rangle_{\boldsymbol{r}} | \tilde{\chi}_{k'\beta'}^{(\nu)} \rangle_{\boldsymbol{R}}] \right] . \tag{14.38}$$

The time-dependent dipole moment relies on a proper description of both intra- and interstate interactions between the TBFs in the underlying simulation, and therefore offers a rather challenging expectation value to simulate. XFAIMS provides a qualitative description of the time-dependent dipole moment oscillations (blue line on Figure 14.7, right panel), despite the use of the critical IFGA (as mentioned in Section 14.5.1, the IFGA was proposed for the simulation of high-dimensional systems) and the SPA0. The dephasing observed at a later time can be corrected by relaxing the IFGA [44]. If the SPA0 is relaxed too (leading to XFFMS, left panel Figure 14.7), the time-dependent moment dipole is in almost perfect agreement with the numerically-exact quantum dynamics (QD) reference (black line on Figure 14.7, right panel). The small dephasing at the end of the simulation is due to the rather limited number of TBFs employed in the FMS simulation, 14 in this particular case. The QD result is perfectly recovered by adding an important number of fixed Gaussian functions in the simulation (orange line on Figure 14.7, right panel). We note that the behavior of the time-dependent dipole moment at a later time, when the nuclear wave packet in S_1 comes back into the Franck–Condon region, is still in qualitative agreement with the QD simulation when using XFAIMS or XFFMS (see Ref. [44]).

While these tests are limited to a single molecular system, they are designed to stress the AIMS approximations in different conditions and provide a satisfactory assessment of their quality. We finally note that other strategies to improve the quality of the SPA0 approximation [37, 38] or efficiently compute Hamiltonian matrix elements have been proposed in the literature [2].

14.5.3 On-the-Fly *Ab Initio* Multiple Spawning

Owing to the SPA0, AIMS dynamics can be performed on-the-fly, i.e., without the need to precompute potential energy surfaces and non-adiabatic couplings before performing the nuclear dynamics. In AIMS, TBFs are propagated using *ab initio* molecular dynamics, that is, by computing the electronic energy and the nuclear gradient at each nuclear integration time step to solve the classical equations of motion (Eqs. (14.14), (14.15), and (14.16)). Then, by computing electronic structure quantities – electronic energies, non-adiabatic coupling vectors, and also spin–orbit coupling or (transition) dipole moments if required – at the centroid positions between all TBFs present in the dynamics (and, if needed, at the position of the TBFs), we can construct the Hamiltonian matrix from the matrix elements in Eq. (14.37) and solve the equations of motion (14.18) (within the SPA0) to propagate the complex coefficients on the support of the TBFs. Formally, the number of electronic-structure calculations required per integration time step, considering a total number of N_{TBFs}, is $\frac{N_{TBFs}\times(N_{TBFs}+1)}{2}$. This computational overhead is the price we have to pay for preserving the couplings between TBFs and not enforcing an independent trajectory approximation as done for example in surface hopping (see Section 14.5.4). While the calculation of Hamiltonian matrix

elements in AIMS can lead to computationally expensive runs when a large number of TBFs are spawned, it is clear from the form of Eq. (14.37) that one can screen for the important matrix elements to be computed by monitoring the overlaps between TBFs [35].

AIMS can be interfaced with any electronic-structure method that can provide the aforementioned electronic-structure quantity. Non-adiabatic dynamics with AIMS were, for example, performed with state-averaged complete active space self-consistent field (SA-CASSCF), multi-state complete active space perturbation of second order (MS-CASPT2), multi-reference configuration interaction (MRCI), linear-response time-dependent density functional theory (LR-TDDFT), or floating occupation molecular orbital complete active space configuration interaction (FOMO-CASCI). More information about these electronic-structure methods can be found in Chapters 2–10 of this book. More recently, AIMS was combined with the electronic-structure code TeraChem [76–79], which capitalizes on the significant acceleration offered by graphics processing units (GPUs). The combination of AIMS with GPU-accelerated electronic structure methods has allowed the photo-chemistry and photo-physics of larger and/or more complex molecular systems with SA-CASSCF [66], LR-TDDFT [13], or FOMO-CASSCI [56] to be studied.

14.5.4 *Ab Initio* Multiple Spawning versus Trajectory Surface Hopping

We briefly summarize here the main differences between AIMS and Trajectory Surface Hopping (TSH, see Chapter 16 for an extensive discussion of this method), which is one of the most commonly employed on-the-fly non-adiabatic dynamics schemes for molecules. In short, TSH portrays the non-adiabatic dynamics of nuclear wave packets as a swarm of independent classical trajectories that can *hop* between electronic adiabatic states. The most commonly employed hopping algorithm – the *Fewest Switches* defined by Tully in 1990 [75] – computes after each nuclear time step a hopping probability (based on the strength of non-adiabatic couplings as well as complex amplitudes propagated along the classical trajectory) that will be used in a stochastic algorithm to determine if a trajectory has to jump from one electronic state to another. TSH and AIMS make use of classical trajectories differently: AIMS propagates Gaussian functions, having a certain width both in position and momentum coordinates, along classical trajectories, while TSH propagates purely classical trajectories, i.e., δ-function like. In AIMS, a Gaussian function is combined with a single complex coefficient and associated with one specific electronic state – any coupling with another electronic state requires the creation of a new TBF in the said state via the spawning algorithm (Section 14.3.2). In TSH, a trajectory carries one complex coefficient for each electronic state considered in the dynamics, and the trajectory can jump onto a different state thanks to the hopping algorithm. In this sense, a TSH trajectory is completely independent of the other trajectories of the swarm, while in AIMS all TBFs of a given parent branch β will be coupled. While the independent classical trajectory approximation (ITA) makes TSH a remarkably simple algorithm to use (all trajectories can be calculated independently), it also leads to the so-called overcoherence problem [8, 10, 14, 18, 69]. Briefly, the overcoherence problem originates from the fact that all the TSH complex coefficients $\tilde{C}^{\alpha}_{\mu}(t)$ associated with a trajectory α are forced to evolve along the same classical trajectory. The equations of motion for the TSH coefficients are given by

$$i\frac{d\tilde{C}^{\alpha}_{\mu}(t)}{dt} = E^{el}_{\mu}(\boldsymbol{R}^{\alpha})\tilde{C}^{\alpha}_{\mu}(t) - i\sum_{\nu}^{\infty}\boldsymbol{d}_{\mu\nu}(\boldsymbol{R}^{\alpha})\cdot\dot{\boldsymbol{R}}^{\alpha}\tilde{C}^{\alpha}_{\nu}(t)\,, \tag{14.39}$$

where $\boldsymbol{R}^{\alpha}$ stands for the position of the trajectory α at time t (and $\dot{\boldsymbol{R}}^{\alpha}$ its classical velocity). Comparing Eq. (14.39) with the equations of motion for the AIMS coefficients (Eq. (14.18) highlights the independent nature of the trajectories in TSH: the TSH coefficients are evolved along the trajectory

α and no couplings with other trajectories are considered (due to the ITA mentioned before). In AIMS, all TBFs from a branch β are coupled together, and a complex coefficient for an electronic state μ is attached to a Gaussian evolving on this particular state only. In a very pictorial way, one could think of a TSH trajectory as a set of infinitely narrow TBFs, one per electronic state, whose dynamics is dictated by a single TBF at the time, the one evolving on the running state. In this picture, all the TBFs are forced to follow the dynamics associated to the TBF on the driving state, until a *hop* takes place, at which point the TBF on the new driving state start driving the dynamics. This picture, while naive, helps us understand one of the central advantages of coupling TBFs in AIMS as compared to the single trajectory picture of TSH: the description of decoherence effects after a region of non-adiabaticity.

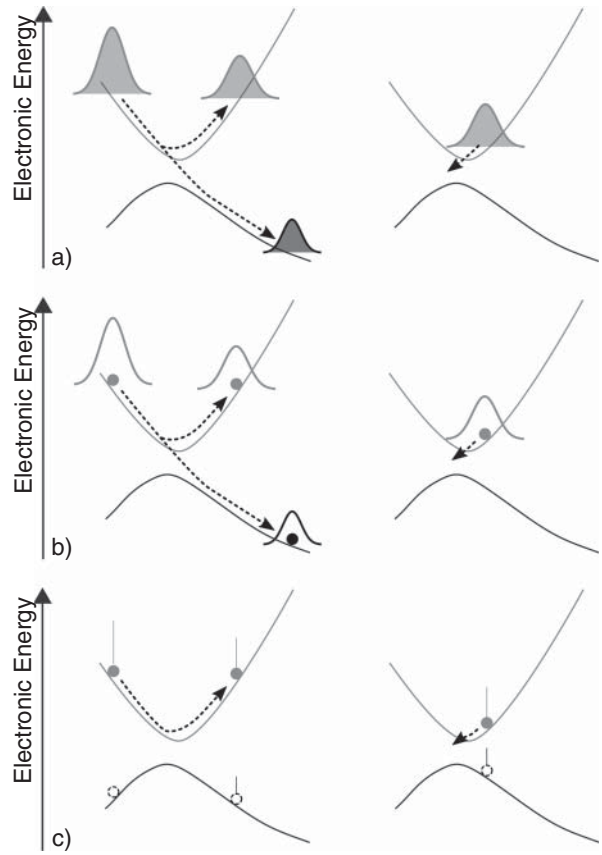


Figure 14.8 Schematic representation of a decoherence event as pictured by exact quantum dynamics (a), AIMS (b), or TSH (c). The first non-adiabatic crossing is depicted on the left panel, while the second passage through the non-adiabatic region is pictured on the right one. In TSH, the trajectory driving the dynamics is indicated by a filled circle, while the "ghost" trajectory on the other state, which follows the driving trajectory, is given by a dashed circle (we consider here that the driving trajectory does not hop and stays on the upper state). The bar on top of the trajectories symbolizes the amount of amplitude carried by each state: 1 for the upper state at the beginning of the dynamics (long bar), shared between both states after the first crossing (smaller bars).

Figure 14.8 gives an example of decoherence following a non-adiabatic event. From a nuclear quantum dynamics perspective (Figure 14.8(a)), a nuclear wave packet enters a non-adiabatic region, where a branching takes place, leaving after the coupling region a substantial amount of the nuclear wave packet on the upper state. Once the crossing is passed, the lower potential energy curve drives the nuclear wave packet (on the lower state) away from the coupling region rapidly (left panel of Figure 14.8(a)). On the other hand, the wave packet on the upper state might reach a repulsive region of the potential energy curve that will push it back towards the intersection region: the two nuclear wave packets *decohere*, i.e., their overlap falls to zero. As the nuclear wave packet on the upper state comes back into the non-adiabatic coupling region (right panel of Figure 14.8(a)), it can transfer additional amplitude to the lower state, but importantly there are in this case *no interferences* between the upper nuclear wave packet and the lower one created in the first crossing event as decoherence happened.

In AIMS (Figure 14.8(b)), a TBF evolving on the upper state reaches the non-adiabatic region, spawns a new TBF in the lower state, and amplitude will be transferred between the two TBFs, which then evolve according to their respective potential energy curve. The TBF on the lower state rapidly leaves the coupling region, while the one on the upper state reaches the repulsive part of the potential and moves back towards the coupling region. As the lower TBF leaves the non-adiabatic coupling region, its coupling with the upper TBF vanishes (the overlap in the last term of the Hamiltonian matrix element between the TBF, Eq. (14.37) and Figure 14.2(a), is close to zero) and no interference between them takes place. In other words, the TBFs have decohered. Hence, the upper TBF, when reaching the coupling region for a second time, spawns a new TBF on the lower state and amplitude transfer takes place.

In TSH (Figure 14.8(c)), a trajectory α is initiated in the upper state, with the complex coefficient corresponding to this state set to 1 and all other coefficients set to zero. When the trajectory reaches the region of strong non-adiabatic coupling, the amplitude is transferred from the upper to the lower state based on Eq. (14.39). The fewest-switches algorithm may allow the trajectory α to remain on the upper electronic state, in which case both the amplitudes on the upper and the lower state are evolved on the support of the α trajectory still evolving on the upper state. When the trajectory reaches the repulsive part of the potential energy curve, it comes back towards the non-adiabatic region. However, the amplitude associated with the lower state still follows the trajectory α and has *not* decohered. When the trajectory reaches the non-adiabatic region, the complex coefficient on the upper part will interact with the one on the lower state, leading to artificial interferences. Hence, TSH would, in this particular case, be unable to describe this non-adiabatic event as a single crossing, unless one artificially suppresses the amplitude on the non-running state – a strategy followed by different methods to include decoherence in TSH [19, 27, 28, 62–64, 70, 71]. It should be noted that the TSH trajectory on the upper state could have jumped during the first passage through the non-adiabatic region, naturally then leaving the coupling region with no decoherence problem. Hence, the overcoherence issue of the original (non-corrected) TSH is particularly critical when multiple crossings between the same electronic states take place. This issue is naturally solved in AIMS by using coupled TBFs, at the cost of relaxing the computationally-interesting ITA of TSH [21, 45]. Recent mixed quantum/classical approaches [10], in particular based on the exact-factorization of the molecular wave function [1, 50, 51], have proposed recovering a detailed description of decoherence effects in non-adiabatic dynamics by coupling trajectories. The interested reader is referred to the Chapter 17 of this book for more information.

14.6 Dissecting an *Ab Initio* Multiple Spawning Dynamics

Section 14.5 presented AIMS from a formal perspective. We propose in this section to explain in more detail the steps involved in an AIMS calculation, to exemplify some of the concepts introduced earlier, and to discuss a practical example of AIMS dynamics.

14.6.1 The Different Steps of an *Ab Initio* Multiple Spawning Dynamics

Different implementations of AIMS have been proposed, in particular of the spawning algorithm, and the flowchart presented in Figure 14.9 presents a simplified AIMS algorithm as included in the software package Molpro [81] (since version 2012) and discussed in great detail in Ref. [35]. Within

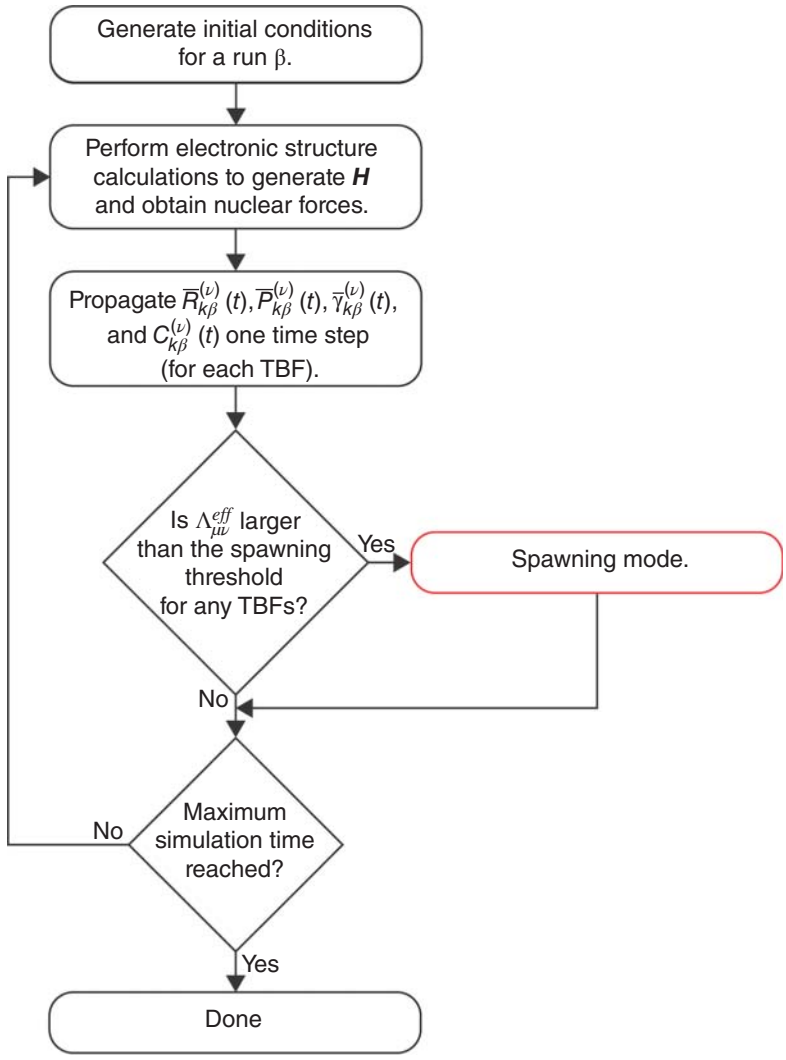


Figure 14.9 Flowchart summarizing the critical steps of a typical algorithm performing *Ab Initio* Multiple Spawning. The Spawning Algorithm (red box) is discussed in more detail in Section 14.3.2. Figure based on Ref. [35].

the IFGA, the AIMS dynamics starts with a single TBF at time $t = 0$ (for a given run β). For each time step, the electronic structure is first calculated for all the TBFs present in the simulation (spawning events keep increasing the number of TBFs) and the Hamiltonian matrix elements (Eq. (14.37), within the SPA0) are constructed. All TBF quantities (positions, momenta, phase, as well as their coupled complex amplitude) are then propagated for one time step. At this point, the spawning algorithm monitors the strength of the effective couplings, and the dynamics can enter a spawning mode and increase the number of TBFs if necessary (see Section 14.3.2 for additional details on the spawning algorithm). Different tests are finally performed to detect potential issues with norm or energy conservation, as well as potential electronic-state flipping during this integration step. If such problems are identified, the time step can be rejected, and its size reduced until tests are eventually passed [35]. If the time step is accepted, the dynamics proceeds further.

14.6.2 Example of *Ab Initio* Multiple Spawning Dynamics – the Photo-Chemistry of Cyclohexadiene

In the following, we discuss how an AIMS simulation takes place in practice and how it can be analyzed, using the non-adiabatic dynamics of cyclohexadiene (CHD) as an example. It should be noted, though, that this section does *not* intend to propose a detailed mechanistic study of the photo-chemistry of this molecule, but only to explain how AIMS is used in practice. Upon light absorption, CHD (see inset of Figure 14.10(b) can be promoted into its first excited electronic state (S_1). This process triggers an excited-state dynamics that can potentially lead to the ring opening of CHD to form hexatriene. The electronic structure employed for the following calculations is

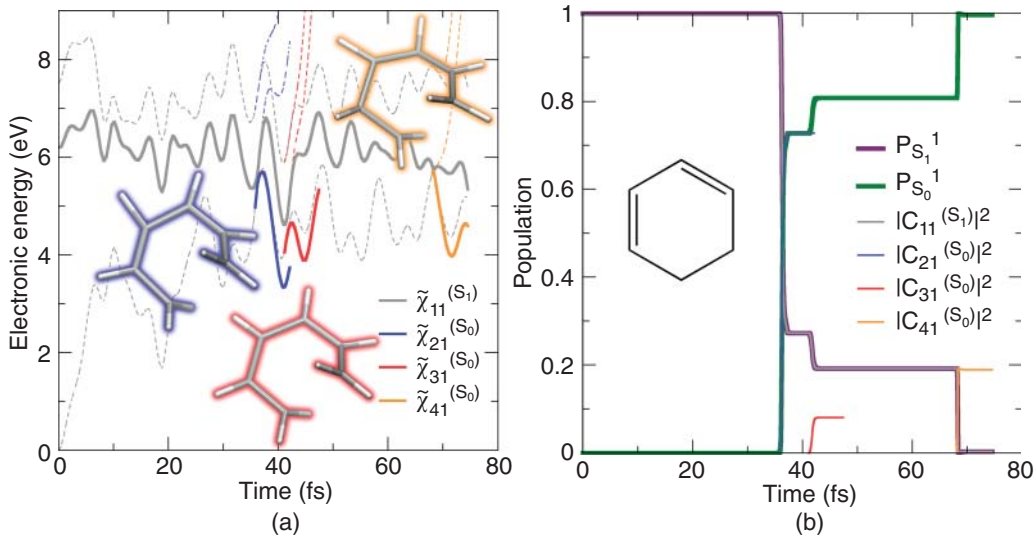


Figure 14.10 Non-adiabatic dynamics of cyclohexadiene with AIMS. Analysis of a single branch ($\beta = 1$) of AIMS dynamics, starting with a TBF in S_1 ($\tilde{\chi}_{11}^{(S_1)}$), which spawns three child TBFs in S_0. (a) Electronic energies along the four TBFs. Full lines indicate the electronic energy followed by each TBF, while the dashed lines (with corresponding colors) highlight the electronic energy of the other electronic states at the current position of the TBF. The three insets depict the molecular geometries at the time of spawning for the three child TBFs. (b) Electronic-state populations during the AIMS dynamics. The S_1 (S_0) population is given by a thick palatinate (green) line, while the TBF populations are given by thin lines with colors matching the ones on panel (a) (see text for more details). The structure of cyclohexadiene is given in the inset.

SA(3)-CASSCF(4/3), combined with a 6-31G* basis set (this level of electronic-structure theory was validated in previous work [30, 31, 66]).

The first step consists of sampling some initial conditions for the TBFs, i.e., determine their initial nuclear positions and momenta. For the most general use of AIMS, we shall consider that the molecule is originally in its ground electronic and vibrational state and will be excited by an ultrashort pulse. In other words, we are simply going to project the ground-state nuclear wave function on the desired electronic state (here S_1), forming a nuclear wave packet at time $t = 0$. More complex initial conditions can be generated, for example when XFAIMS is employed and the photo-excitation process is treated explicitly. A general strategy to sample positions and momenta for a ground vibrational nuclear wave function is to compute its Wigner distribution [55]. For molecular systems, one actually computes an approximate Wigner distribution for a set of uncoupled harmonic oscillators (for more details on this subject, the reader is referred to Refs. [11, 55] as well as Chapter 16). The construction of this approximate Wigner distribution only requires a ground-state optimized geometry and corresponding harmonic frequencies – quantities that can easily be obtained with most electronic-structure methods. A sampling is performed from this distribution, providing for each selected point a set of positions and momenta. These will define the initial conditions, $\bar{\boldsymbol{R}}_{k\beta}^{(\mu)}(t=0)$ and $\bar{\boldsymbol{P}}_{k\beta}^{(\mu)}(t=0)$, for each parent TBF β in electronic state μ (in the case of CHD, μ corresponds to S_1). Within the IFGA, all parent TBFs will be run independently, meaning that the complex amplitude for the parent TBF k in branch β attributed to state μ is simply set to unity at the beginning of the dynamics: $C_{k\beta}^{(\mu)}(t=0) = (1.0, 0.0)$. The AIMS dynamics is then initiated with only the parent TBF present until it encounters a region of strong non-adiabaticity, where the spawning of a new child TBF can take place. The dynamics continues, with new spawning events possible for both parent and child TBFs, until a certain termination criterion is met, for example when most of the nuclear wave packet is returned to the ground electronic state.

Figure 14.10 gives an example of the dynamics of a parent TBF and the creation of child TBFs for the excited-state dynamics of CHD. As mentioned above, only the parent TBF is present at time $t = 0$. The parent has the label $k = 1$ (being the first TBF of the branch), is on branch $\beta = 1$, and is running on the electronic state S_1: its notation is therefore $\tilde{\chi}_{11}^{(S_1)}$. The electronic energy of the parent TBF during the AIMS run is indicated by the thick gray line in Figure 14.10(a). The corresponding time-trace for the population of this TBF, $|C_{11}^{(S_1)}(t)|^2$, is given by a similar line in Figure 14.10(b). The parent TBF evolves originally adiabatically, i.e., without any strong interaction with the other electronic states. However, after 30 fs of dynamics, the S_0 state comes closer (in energy) to the running state S_1 (lower dashed gray line in Figure 14.10(a)), until both states become nearly degenerate (at $t \sim 36$ fs). At this point, the parent TBF enters a region of strong coupling and spawns a new TBF, $\tilde{\chi}_{12}^{(S_0)}$, on the coupled state ($\beta = 1$, $k = 2$, $\mu = S_0$), depicted by a thick blue line. The two lines become nearly degenerate at the time of the spawning, and it is interesting to note that the blue segment before the intersection region results from the dynamics from t_{entry} to t_{spawn}, as detailed in Section 14.3.2. The spawning geometry for $\tilde{\chi}_{12}^{(S_0)}$ is given as an inset in Figure 14.10(a) (with a blue shadow). The child TBF $\tilde{\chi}_{12}^{(S_0)}$ evolves on the ground electronic state S_0 and relaxes towards a lower electronic energy rapidly, pushing the excited electronic states higher in energy (see dashed blue lines in Figure 14.10(a)). When looking at the transfer of amplitude, the parent TBF efficiently transfers a large amount of its population to the first child TBF (large variations in the gray and blue curves in Figure 14.10(b) at around 36 fs). When the TBF leaves the coupling region and is no more coupled to any other TBFs, its dynamics can be stopped (explaining why the curves for $\tilde{\chi}_{12}^{(S_0)}$ stops after 42 fs). The parent TBF continues its evolution in S_1 and rapidly hits again the S_1/S_0 intersection seam, spawning a second child TBF, $\tilde{\chi}_{13}^{(S_0)}$ (red lines in Figure 14.10), with whom it only exchanges a small amount of population ($\sim 8\%$). After 68 fs of dynamics, the parent TBF

finally hits for the last time the intersection seam, transferring almost all its remaining amplitude to the newly-created TBF $\tilde{\chi}_{14}^{(S_0)}$ (orange). This AIMS run highlights the interaction between TBFs and the fact that they all have their independent dynamics while being connected via their complex coefficients.

In the previous paragraph, we discussed the concept of "TBF population", given by $|C_{k\beta}^{(\mu)}(t)|^2$. It is important to realize that the population of a given electronic state, say λ, is *not* obtained by simply adding up all the TBF populations for TBF evolving in λ, as we are using a non-orthogonal Gaussian basis set. The population in state λ for a given branch β is obtained by calculating the expectation value of the projector $\hat{\mathcal{P}}_\lambda = |\Phi_\lambda\rangle\langle\Phi_\lambda|$ using the AIMS molecular wave function:

$$\begin{aligned} P_\lambda^\beta(t) &= \sum_{\mu\nu}^{\infty}\left[\sum_{k}^{N_\mu^\beta(t)}\sum_{k'}^{N_\nu^\beta(t)}(C_{k\beta}^{(\mu)}(t))^* C_{k'\beta}^{(\nu)}(t)\langle\Phi_\mu\tilde{\chi}_{k\beta}^{(\mu)}|\hat{\mathcal{P}}_\lambda|\tilde{\chi}_{k'\beta}^{(\nu)}\Phi_\nu\rangle_{\boldsymbol{r},\boldsymbol{R}}\right] \\ &= \sum_{kk'}^{N_\lambda^\beta(t)}(C_{k\beta}^{(\lambda)}(t))^* C_{k'\beta}^{(\lambda)}(t)\langle\tilde{\chi}_{k\beta}^{(\lambda)}|\tilde{\chi}_{k'\beta}^{(\lambda)}\rangle_{\boldsymbol{R}} = \sum_{kk'}^{N_\lambda^\beta(t)}(C_{k\beta}^{(\lambda)}(t))^* C_{k'\beta}^{(\lambda)}(t)(\boldsymbol{S})_{k\beta,k'\beta}^{\lambda\lambda}\,. \end{aligned} \tag{14.40}$$

Equation (14.40) highlights the potential importance of the Gaussian interference terms in the calculation of AIMS expectation values. The population trace for the S_1 and S_0 states of the branch $\beta = 1$ ($P_\lambda^1(t)$, with $\lambda = S_1, S_0$) is given in Figure 14.10(b) with thick lines.

The results presented up to now were all for a single branch $\beta = 1$, i.e., the AIMS dynamics originating from a single parent TBF (a single initial condition). We know, however, that we need more than one branch to reproduce the dynamics of the initial nuclear wave packet adequately. How should we then compute expectation values in AIMS?

In FMS, one would compute the expectation value of a given operator $\hat{\mathcal{O}}$ as

$$\begin{aligned} \mathcal{O}(t) &= \frac{\langle\Psi(t)|\hat{\mathcal{O}}|\Psi(t)\rangle_{\boldsymbol{r},\boldsymbol{R}}}{\langle\Psi(t)|\Psi(t)\rangle_{\boldsymbol{r},\boldsymbol{R}}} \\ &= \frac{\sum_{\beta\beta'}^{N_{ini}}\sum_{\mu\nu}^{\infty}\sum_{k}^{N_\mu^\beta(t)}\sum_{k'}^{N_\nu^{\beta'}(t)}(C_{k\beta}^{(\mu)}(t))^* C_{k'\beta'}^{(\nu)}(t)\langle\Phi_\mu\tilde{\chi}_{k\beta}^{(\mu)}|\hat{\mathcal{O}}|\tilde{\chi}_{k'\beta'}^{(\nu)}\Phi_\nu\rangle_{\boldsymbol{r},\boldsymbol{R}}}{\sum_{\beta\beta'}^{N_{ini}}\sum_{\mu\nu}^{\infty}\sum_{k}^{N_\mu^\beta(t)}\sum_{k'}^{N_\nu^{\beta'}(t)}(C_{k\beta}^{(\mu)}(t))^* C_{k'\beta'}^{(\nu)}(t)(\boldsymbol{S})_{k\beta,k'\beta'}^{\mu\nu}}\,, \end{aligned} \tag{14.41}$$

where we used the definition of the FMS molecular wave function Eq. (14.17) and considered the spawning algorithm (the number of TBFs is time-dependent). If one now applies the IFGA, Eq. (14.41) reduces to

$$\mathcal{O}(t) \approx \frac{1}{N_{ini}}\sum_{\beta}^{N_{ini}}\left[\frac{\sum_{\mu\nu}^{\infty}\sum_{k}^{N_\mu^\beta(t)}\sum_{k'}^{N_\nu^\beta(t)}(C_{k\beta}^{(\mu)}(t))^* C_{k'\beta}^{(\nu)}(t)\langle\Phi_\mu\tilde{\chi}_{k\beta}^{(\mu)}|\hat{\mathcal{O}}|\tilde{\chi}_{k'\beta}^{(\nu)}\Phi_\nu\rangle_{\boldsymbol{r},\boldsymbol{R}}}{\sum_{\mu\nu}^{\infty}\sum_{k}^{N_\mu^\beta(t)}\sum_{k'}^{N_\nu^\beta(t)}(C_{k\beta}^{(\mu)}(t))^* C_{k'\beta}^{(\nu)}(t)(\boldsymbol{S})_{k\beta,k'\beta}^{\mu\nu}}\right], \tag{14.42}$$

where interferences between TBFs coming from different branches β are neglected and the final expectation value is obtained by averaging *incoherently* over all the branches (or initial conditions). Hence, replacing $\hat{\mathcal{O}}$ by $\hat{\mathcal{P}}_\lambda = |\Phi_\lambda\rangle\langle\Phi_\lambda|$ in Eq. (14.42) gives,

$$P_\lambda(t) \approx \frac{1}{N_{ini}}\sum_{\beta}^{N_{ini}}\left[\frac{\sum_{kk'}^{N_\lambda^\beta(t)}(C_{k\beta}^{(\lambda)}(t))^* C_{k'\beta}^{(\lambda)}(t)(\boldsymbol{S})_{k\beta,k'\beta}^{\lambda\lambda}}{\sum_{\mu\nu}^{\infty}\sum_{k}^{N_\mu^\beta(t)}\sum_{k'}^{N_\nu^\beta(t)}(C_{k\beta}^{(\mu)}(t))^* C_{k'\beta}^{(\nu)}(t)(\boldsymbol{S})_{k\beta,k'\beta}^{\mu\nu}}\right], \tag{14.43}$$

which, considering that the total (FMS or AIMS) molecular wave function is normalized, corresponds to an average of Eq. (14.40) over all N_{ini} branches. In other words, the AIMS expectation

value for the population of a given electronic state is obtained by propagating several initial conditions (several branches) and averaging them together. More information about the calculation of expectation values in AIMS and FMS can be found in Ref. [11].

Returning to the non-adiabatic dynamics of CHD, one needs to run more AIMS branches to perform the average given in Eq. (14.43). A converged AIMS usually consists of tens to hundreds of independent branches, depending on the type of expectation values or quantities one wants to calculate. As an example, we summarize here the results obtained with 12 different branches ($N_{ini} = 12$) for CHD (we should keep in mind that more initial conditions would be desirable to converge the simulation fully). Figure 14.11(a) presents the time trace for the population of state S_1 ($P_{S_1}(t)$, as defined in Eq. (14.43)), averaged over 12 branches whose initial configurations are given in the inset. The nuclear wave packet, originally generated in S_1, relaxes towards the ground electronic state in less than 180 fs. The S_1 population for each individual branch ($P^{\beta}_{S_1}(t)$, as obtained from Eq. (14.40)) is reported as dashed lines in Figure 14.11(a). While the AIMS dynamics was initiated with 12 parent TBFs, the spawning algorithm allowed for an important increase in the number of TBFs, reaching 72 by the end of this non-adiabatic dynamics.

A simple and informative analysis that can be performed on an AIMS simulation is the projection of the TBFs onto a specific molecular coordinate. In the particular case of CHD, we focus on the C—C bond (see inset of Figure 14.11(a)) responsible for the ring opening. Figure 14.11(b) shows a projection of the 72 TBFs generated during the AIMS dynamics onto this coordinate, the thickness of each line being proportional to the TBF population $|C^{(\mu)}_{k\beta}|^2$. Each line represents a TBF, and is depicted in palatinate if the TBF evolves in S_1, in green for S_0, and orange for S_2. This representation clearly shows the spreading of the AIMS wave packets during the non-adiabatic dynamics. At the

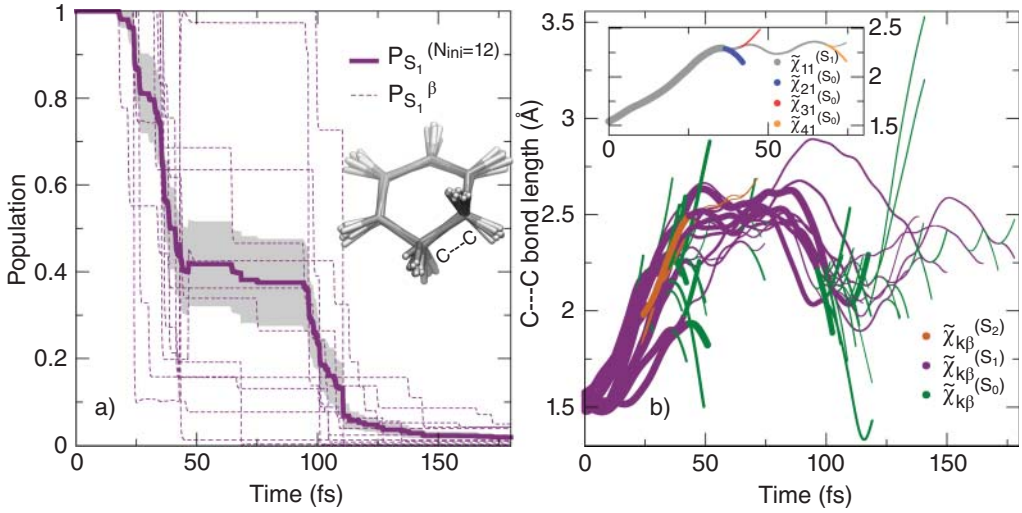


Figure 14.11 Non-adiabatic dynamics of cyclohexadiene with AIMS. Analysis of the results obtained using 12 initial conditions (or branches), generating 72 TBFs in total. (a) The overall population of the S_1 state, $P_{S_1}(t)$, obtained by averaging the 12 $P^{\beta}_{S_1}(t)$, is given by a thick palatinate line, while the 12 individual S_1 populations, $P^{\beta}_{S_1}(t)$, are given by dashed lines. The gray area indicates the standard error of the AIMS population. The inset shows the configuration of the 12 initial conditions. (b) Projection of all the TBFs propagated in AIMS onto the ring-opening coordinate, C—C. The thickness of each line is proportional to the TBF population. Palatinate lines depict TBFs in S_1, green lines those in S_0, and orange lines those in S_2. The inset highlights the same projection, but only for the branch $\beta = 1$ presented in Figure 14.10 using the same color code.

beginning of the dynamics, when the nuclear wave packet formed in S_1 is in the Franck–Condon region, the TBFs exhibit standard C—C bond lengths (around 1.54 Å). However, the wave packet rapidly starts to leave the Franck–Condon region (stretch of the C—C bond), and approaches a non-adiabatic region after only 25 fs of dynamics, leading to the spawn of new TBFs in S_0 (green lines appear in Figure 14.11(b)) and population transfer towards S_0 (green lines become thicker, and population in S_1 decays, as also observed in Figure 14.11(a)). The TBFs in S_1 keep evolving with an elongated C—C, until they reach the intersection seam again and transfer population again towards the ground electronic state. We note that the TBFs in S_0, upon creation, show a rapid C—C stretch (towards a ring opening) or C—C contraction (towards a ring closure).

The inset of Figure 14.11(b) shows the same projection, but only for the branch $\beta = 1$ described above in Figure 14.10 (the color code in the inset is chosen to match the one introduced in the previous discussion of the branch $\beta = 1$). The inset shows that the C—C bond length of the parent TBF, $\tilde{\chi}_{11}^{(S_1)}$, rapidly stretches and hits the intersection seam, leading to the spawning of a new TBF, $\tilde{\chi}_{21}^{(S_0)}$, and to amplitude transfer (the gray line becomes thinner and the blue one thicker – see also Figure 14.10(b)). The C—C bond length of $\tilde{\chi}_{21}^{(S_0)}$ appears to evolve towards ring closure, following the dynamics of the parent TBF at the time of the spawning. The parent TBF spawns a second time, forming $\tilde{\chi}_{31}^{(S_0)}$ (red line in the inset of Figure 14.11(b)) and leading to a smaller population transfer. Interestingly though, this spawn takes place when the C—C bond of $\tilde{\chi}_{11}^{(S_1)}$ is in an elongation process, leading to a child TBF, $\tilde{\chi}_{31}^{(S_0)}$, evolving towards a ring opening. The parent TBF finally spawns a last TBF, transferring most of its remaining amplitude (as described in the analysis above).

14.7 *In Silico* Photo-Chemistry with *Ab Initio* Multiple Spawning

Over the last decades, AIMS has helped shed light on the non-adiabatic dynamics of a wide range of organic as well as inorganic molecules (a table summarizing different molecular applications of AIMS is available in Ref. [11]). In the following, we describe representative applications of AIMS to the non-adiabatic dynamics of three different molecules, demonstrating the potential and predictive power of this method.

Ethylene – C_2H_4 (inset of Figure 14.12) – may appear to be a small and simple molecule, but it suffers a surprisingly complex excited-state dynamics and is considered as a prototypical molecule

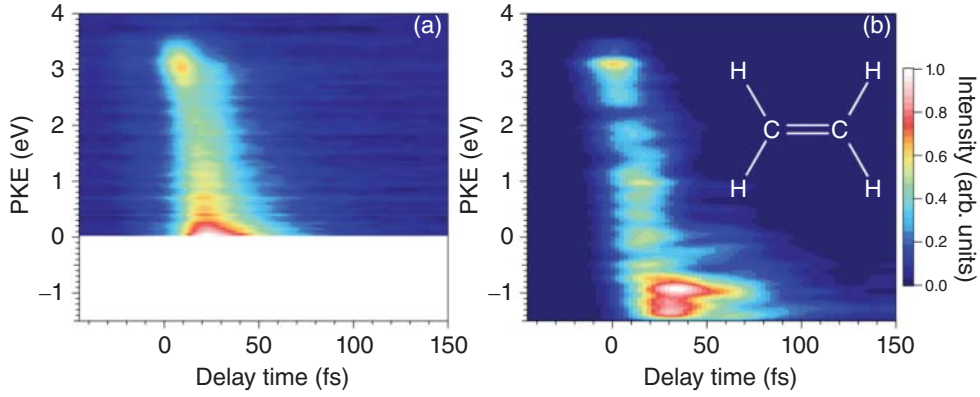


Figure 14.12 (a) Experimental and (b) predicted (AIMS/MS-CASPT2) time-resolved photo-electron kinetic energy spectra of ethylene (inset). Adapted from Ref. [32]. Copyright 2015 American Chemical Society.

to study photo-isomerization around a double bond. As such, ethylene has been the subject of a number of important experimental and theoretical studies. AIMS combined with MS-CASPT2 [74] has been employed to study the non-adiabatic dynamics following photo-excitation of ethylene in its bright $\pi\pi^*$ excited state [73]. An excellent agreement between experiment and AIMS was observed, in particular when comparing an AIMS-simulated femtosecond time-resolved photo-electron spectrum with the measured total ion yield from a vacuum UV/vacuum UV pump–probe experiment [73]. This analysis further explained why the theoretical S_1 lifetimes predicted by different excited-state dynamics methods are significantly longer (89–180 fs) than the experimentally observed lifetime (~50 fs): the photo-excited ethylene molecule can enter into a region of the S_1 potential energy surface where the experimental probe pulse is insufficiently energetic to induce an ionization process. Therefore, the decay of the measured photo-ion yield does not directly correspond to the decay of the S_1 population (when using such a probe pulse), explaining the mismatch between experiment and theory. Another validation of the accuracy of AIMS/MS-CASPT2 can be found in Figure 14.12, which compares the experimental (a) and AIMS-simulated (b) time-resolved photo-electron kinetic energy spectra of ethylene. The spectra obtained with AIMS [52] largely reproduce the trends observed experimentally [32], in particular, the fast decrease of the photo-electron kinetic energy. It is important to note that the AIMS/MS-CASPT2 spectra reported in Figure 14.12(b) was in fact predicted three years before the experimental spectra in Figure 14.12(a) could be recorded. More recently, AIMS was used to simulate a time-resolved x-ray absorption spectroscopy (TRXAS) experiment performed on ethylene, predicting that this technique could potentially provide a signature of the dynamics near conical intersections [53]. Interestingly, such dynamics near conical intersections has recently been observed for the molecule CF_3I using ultrafast gas-phase electron diffraction combined with AIMS dynamics, which closely reproduce the experimental observables and support the experimental interpretation [85].

As described in Section 14.5.3, the development of GPU-accelerated electronic-structure methods paved the way for the non-adiabatic dynamics of even larger (or more complex) molecular systems. The combination of GPU-accelerated SA-CASSCF [26, 67, 68] with AIMS permits large molecular systems to be treated while offering an accurate description of both the electronic structure and the nuclear dynamics – in other words, it allows the compromise between efficiency and accuracy inherent to any non-adiabatic molecular dynamics simulation to be pushed further. The potential of this combination is particularly well illustrated by a recent study focusing on the non-radiative deactivation of the provitamin-D_3 molecule [66]. Provitamin-D_3 contains a cyclohexadiene moiety (see inset of Figure 14.13) that, as described in Section 14.6, can absorb light and undergo a ring-opening mechanism in the excited state. Upon non-radiative relaxation to the ground electronic state, either the parent provitamin-D_3 is recovered or a ring-open form, called previtamin-D_3, can be observed. Further reorganization of previtamin-D_3 leads to the well-known vitamin D molecule. The photo-chemistry of provitamin-D_3 has been studied using AIMS with GPU-SA-CASSCF(6/4) for the electronic structure[7], considering the first three electronic states: the ground state S_0, the first excited state S_1 (bright state), as well as a the second-state S_2 (exhibiting a double-excitation character). The molecule considered has 51 atoms, and up to 330 TBFs in total are generated during the 2 ps of non-adiabatic dynamics. Experimentally, the non-radiative decay of provitamin-D_3 exhibits a bi-exponential decay – a behavior closely reproduced by the AIMS/GPU-SA-CASSCF non-adiabatic (Figure 14.13). Fitting the decay of the S_1 population

7 This level of theory was benchmarked against high-level calculations for the cyclohexadiene/hexatriene molecules.

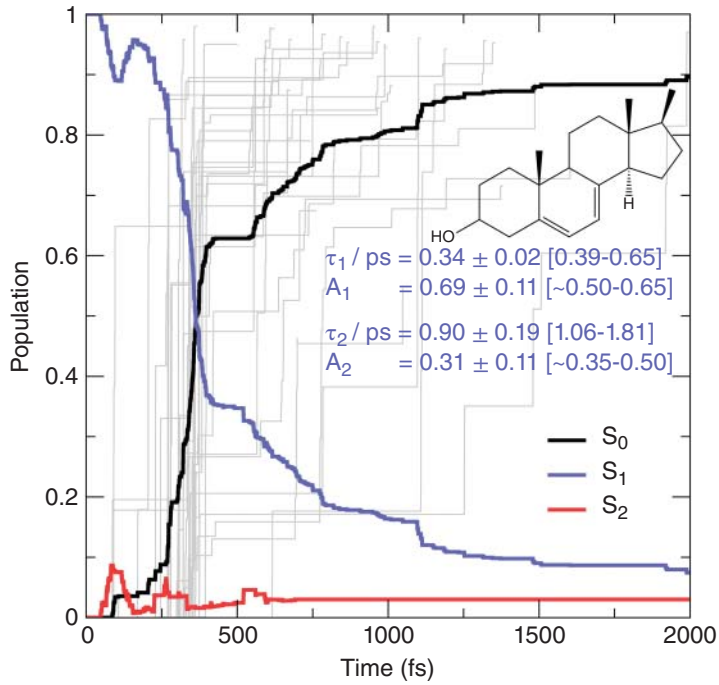


Figure 14.13 Time trace of the electronic population of provitamin-D_3 upon photo-excitation in S_1. The thick lines give the average AIMS populations (S_0 in black, S_1 in blue, and S_2 in red), while the gray lines indicate the S_0 population for each individual AIMS run. Reprinted from Ref. [66]. Copyright 2016 American Chemical Society.

provides lifetimes and amplitudes in close agreement with experimental observations (inset of Figure 14.13). Further investigations of the AIMS dynamics revealed that the bi-exponential decay is caused by a non-equilibrium dynamics upon photo-excitation linked to the ring opening and closing motion. This theoretical interpretation of the bi-exponential decay was later validated by experiments based on ultrafast transient absorption spectroscopy [65].

The chemistry taking place upon photo-excitation of the small molecule thioformaldehyde-S-oxide (see Figure 14.14(b), called in the following "sulfine") is surprisingly rich. Experimentally, a large number of somewhat exotic molecules could be observed upon irradiation of sulfine in an argon matrix [59] (see molecules depicted in Figure 14.14(c)). Are all these molecules formed in the excited states, or are they the result of a subsequent hot ground-state dynamics? AIMS/MS-CASPT2, combined with *ab initio* molecular dynamics using GPU-accelerated density functional theory (DFT), has been employed to answer this question [48]. Upon photo-excitation, the sulfine molecule relaxes towards the ground electronic state in less than 1 ps, coming back to its original form or generating the cyclic molecule oxathiirane, the latter case happening for $\sim 30\%$ of the original sulfine population (Figure 14.14(a)). By further propagating all the TBFs formed in the ground electronic state using GPU-accelerated *ab initio* molecular dynamics, all the molecules detected experimentally could subsequently be observed in the simulation (Figure 14.14(c)). Hence, the non-adiabatic dynamics of the sulfine only leads to the formation of oxathiirane (or regeneration of the original sulfine), and the subsequent hot dynamics in the ground state allows for the generation of the other molecules. However, the timescale for the formation of these molecules in S_0 is much shorter than one would expect from transition state theory, which implies

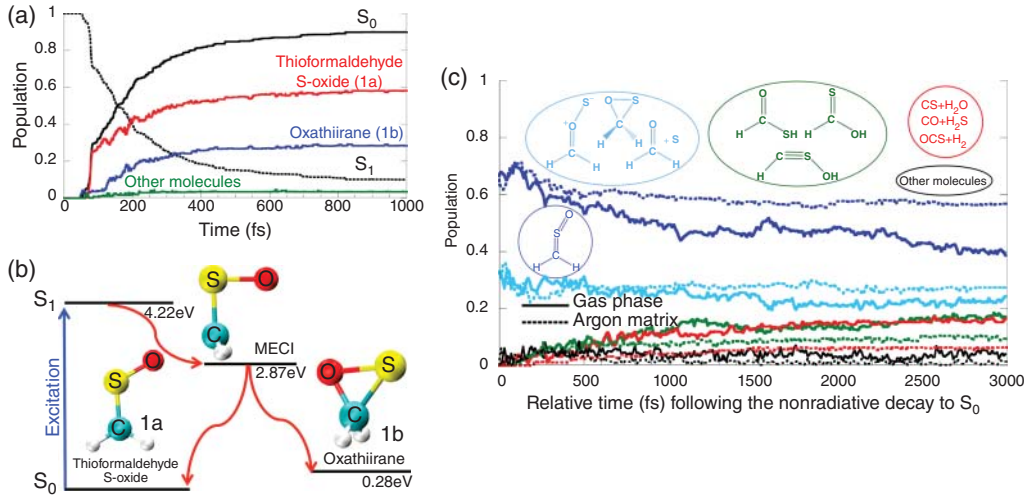


Figure 14.14 Non-adiabatic and hot ground state dynamics of the sulfine molecule. (a) Electronic-state population and molecules observed during the AIMS/MS-CASPT2 dynamics. (b) Critical points of the sulfine potential energy surfaces connecting the Franck–Condon geometry with the minimum energy conical intersection (MECI) and the oxathiirane ground-state geometry. (c) Population of the different molecules formed in the ground-state *ab initio* molecular dynamics following the relaxation of the wave packet in the ground state. For clarity, the different molecules formed are grouped in families (see structures), and a corresponding color code is used to plot the populations. Reprinted from Ref. [48] with permission from Wiley.

that a non-statistical dynamics takes place in the ground electronic state [48]. In other words, the molecule does not have the time to vibrationally relax upon non-adiabatic relaxation, and the large amount of kinetic energy released after the passage through the conical intersection triggers ground-state chemical reactions in a non-statistical (or athermal) way.

14.8 Summary

This chapter introduced the framework of FMS – a method for non-adiabatic quantum molecular dynamics that portrays nuclear wave functions as adaptive linear combinations of frozen Gaussians following classical trajectories. In the limit of a large number of trajectory basis functions and an exact evaluation of their couplings, the FMS framework becomes, in principle, exact. Approximations have been devised around the FMS strategy to allow for the on-the-fly non-adiabatic dynamics of molecules in their full configuration space – a method coined *Ab Initio* Multiple Spawning. *Ab Initio* Multiple Spawning naturally preserves a proper description of decoherence effects following non-adiabatic transitions, thanks to the coupling between trajectory basis functions. Owing to the straightforward derivation of their running equations from first principles, both FMS and *Ab Initio* Multiple Spawning can easily be extended to the description of intersystem crossings or the coupling with an external time-dependent electric field. We detailed the different steps of an *Ab Initio* Multiple Spawning dynamics for the excited-state dynamics of the cyclohexadiene molecule, and summarized some successful applications of this method to molecular systems of experimental interest.

References

1 Agostini, F., Min, S.K., Abedi, A., and Gross, E.K.U. (2016). Quantum-classical non-adiabatic dynamics: coupled- vs. independent-trajectory methods. *J. Chem. Theory Comput.* 12 (5): 2127–2143.

2 Alborzpour, J.P., Tew, D.P., and Habershon, S. (2016). Efficient and accurate evaluation of potential energy matrix elements for quantum dynamics using Gaussian process regression. *J. Chem. Phys.* 145 (17): 174112.

3 Ben-Nun, M. and Martínez, T.J. (1998). Nonadiabatic molecular dynamics: validation of the multiple spawning method for a multidimensional problem. *J. Chem. Phys.* 108: 7244–7257.

4 Ben-Nun, M. and Martínez, T.J. (2000). A multiple spawning approach to tunneling dynamics. *J. Chem. Phys.* 112 (14): 6113–6121.

5 Ben-Nun, M. and Martínez, T.J. (2002). Ab initio quantum molecular dynamics. *Adv. Chem. Phys.* 121: 439–512.

6 Ben-Nun, M. and Martínez, T.J. (2007). A continuous spawning method for non-adiabatic dynamics and validation for the zero-temperature spin-boson problem. *Isr. J. Chem.* 47 (1): 75–88.

7 Ben-Nun, M., Quenneville, J., and Martínez, T.J. (2000). Ab initio multiple spawning: photochemistry from first principles quantum molecular dynamics. *J. Phys. Chem. A* 104: 5161–5175.

8 Bittner, E.R. and Rossky, P.J. (1995). Quantum decoherence in mixed quantum-classical systems: nonadiabatic processes. *J. Chem. Phys.* 103: 8130.

9 Born, M. and Huang, K. (1954). *Dynamical Theory of Crystal Lattices*. Clarendon, Oxford.

10 Crespo-Otero, R. and Barbatti, M. (2018). Recent advances and perspectives on non-adiabatic mixed quantum–classical dynamics. *Chem. Rev.* 118 (15): 7026–7068.

11 Curchod, B.F.E. and Martínez, T.J. (2018). Ab initio nonadiabatic quantum molecular dynamics. *Chem. Rev.* 118 (7): 3305–3336.

12 Curchod, B.F.E., Rauer, C., Marquetand, P. et al. (2016). Communication: GAIMS–generalized ab initio multiple spawning for both internal conversion and intersystem crossing processes. *J. Chem. Phys.* 144 (10): 101102.

13 Curchod, B.F.E., Sisto, A., and Martínez, T.J. (2017). Ab initio multiple spawning photochemical dynamics of DMABN using GPUs. *J. Phys. Chem. A* 121 (1): 265–276.

14 Curchod, B.F.E. and Tavernelli, I. (2013). On trajectory-based nonadiabatic dynamics: Bohmian dynamics versus trajectory surface hopping. *J. Chem. Phys.* 138: 184112.

15 Domcke, W., Yarkony, D., and Köppel, H. (2004). *Conical Intersections: Electronic Structure, Dynamics & Spectroscopy*, vol. 15. World Scientific Pub Co Inc.

16 Fedorov, D.A., Lykhin, A.O., and Varganov, S.A. (2018). Predicting intersystem crossing rates with aims-dft molecular dynamics. *J. Phys. Chem. A* 122 (13): 3480–3488.

17 Fedorov, D.A., Pruitt, S.R., Keipert, K. et al. (2016). Ab initio multiple spawning method for intersystem crossing dynamics: spin-forbidden transitions between $^3B_1$ and $^1A_1$ states of GeH_2. *J. Phys. Chem. A* 120 (18): 2911–2919.

18 Gao, X. and Thiel, W. (2017). Non-hermitian surface hopping. *Phys. Rev. E* 95: 013308.

19 Granucci, G. and Persico, M. (2007). Critical appraisal of the fewest switches algorithm for surface hopping. *J. Chem. Phys.* 126: 134114.

20 Granucci, G., Persico, M., and Spighi, G. (2012). Surface hopping trajectory simulations with spin–orbit and dynamical couplings. *J. Chem. Phys.* 137 (22): 22A501.

21 Granucci, G., Persico, M., and Zoccante, A. (2010). Including quantum decoherence in surface hopping. *J. Chem. Phys.* 133: 134111.

22 Hack, M.D., Wensmann, A.M., Truhlar, D.G. et al. (2001). Comparison of full multiple spawning, trajectory surface hopping, and converged quantum mechanics for electronically non-adiabatic dynamics. *J. Chem. Phys.* 115: 1172.

23 Heller, E.J. (1975). Time-dependent approach to semiclassical dynamics. *J. Chem. Phys.* 62: 1544–1555.

24 Heller, E.J. (1981). Frozen Gaussians: a very simple semiclassical approximation. *J. Chem. Phys.* 75 (6): 2923–2931.

25 Heller, E.J. (1981). The semiclassical way to molecular spectroscopy. *Acc. Chem. Res.* 14 (12): 368–375.

26 Hohenstein, E.G., Luehr, N., Ufimtsev, I.S., and Martínez, T.J. (2015). An atomic orbital-based formulation of the complete active space self-consistent field method on graphical processing units. *J. Chem. Phys.* 142 (22): 224103.

27 Jaeger, H.M., Fischer, S., and Prezhdo, O.V. (2012). Decoherence-induced surface hopping. *J. Chem. Phys.* 137: 22A545.

28 Jasper, A.W. and Truhlar, D.G. (2007). Electronic decoherence time for non-born-oppenheimer trajectories. *J. Chem. Phys.* 127: 194306.

29 Joubert-Doriol, L., Sivasubramanium, J., Ryabinkin, I.G., and Izmaylov, A.F. (2017). Topologically correct quantum non-adiabatic formalism for on-the-fly dynamics. *J. Phys. Chem. Lett.* 8 (2): 452–456.

30 Kim, J., Tao, H., Martinez, T.J., and Bucksbaum, P. (2015). Ab initio multiple spawning on laser-dressed states: a study of 1, 3-cyclohexadiene photoisomerization via light-induced conical intersections. *J. Phys. B Atomic Mol. Phys.* 48 (16): 164003.

31 Kim, J., Tao, H., White, J.L. et al. (2011). Control of 1, 3-cyclohexadiene photoisomerization using light-induced conical intersections. *J. Phys. Chem. A* 116 (11): 2758–2763.

32 Kobayashi, T., Horio, T., and Suzuki, T. (2015). Ultrafast deactivation of the $\pi\pi^*$(V) state of ethylene studied using sub-20 fs time-resolved photoelectron imaging. *J. Phys. Chem. A* 119 (36): 9518–9523.

33 Lasorne, B., Bearpark, M.J., Robb, M.A., and Worth, G.A. (2006). Direct quantum dynamics using variational multi-configuration Gaussian wave packets. *Chem. Phys. Lett.* 432 (4): 604–609.

34 Lasorne, B., Robb, M., and Worth, G.A. (2007). Direct quantum dynamics using variational multi-configuration Gaussian wave packets. Implementation details and test case. *Phys. Chem. Chem. Phys.* 9 (25): 3210–3227.

35 Levine, B.G., Coe, J.D., Virshup, A.M., and Martínez, T.J. (2008). Implementation of ab initio multiple spawning in the Molpro quantum chemistry package. *Chem. Phys.* 347 (1): 3–16.

36 Mai, S., Marquetand, P., and González, L. (2015). A general method to describe intersystem crossing dynamics in trajectory surface hopping. *Int. J. Quantum Chem.* 115 (18): 1215–1231.

37 Makhov, D., Symonds, C., Fernandez-Alberti, S., and Shalashilin, D. (2017). Ab initio quantum direct dynamics simulations of ultrafast photochemistry with multiconfigurational Ehrenfest approach. *Chem. Phys.* 493: 200.

38 Makhov, D.V., Glover, W.J., Martinez, T.J., and Shalashilin, D.V. (2014). Ab initio multiple cloning algorithm for quantum non-adiabatic molecular dynamics. *J. Chem. Phys.* 141 (5): 054110.

39 Marian, C.M. (2012). Spin–orbit coupling and intersystem crossing in molecules. *WIREs Comput. Mol. Sci.* 2: 187–120.

40 Martínez, T.J., Ben-Nun, M., and Levine, R.D. (1996). Multi-electronic-state molecular dynamics: a wave function approach with applications. *J. Phys. Chem.* 100 (19): 7884–7895.

41 Martínez, T.J. and Levine, R.D. (1997). Non-adiabatic molecular dynamics: split-operator multiple spawning with applications to photodissociation. *J. Chem. Soc. Faraday Trans.* 93 (5): 941–947.

42 Meek, G.A. and Levine, B.G. (2016). The best of both reps–diabatized Gaussians on adiabatic surfaces. *J. Chem. Phys.* 145 (18): 184103.

43 Mendive-Tapia, D., Lasorne, B., Worth, G.A. et al. (2012). Towards converging non-adiabatic direct dynamics calculations using frozen-width variational Gaussian product basis functions. *J. Chem. Phys.* 137 (22): 22A548.

44 Mignolet, B. and Curchod, B.F.E. (2018). A walk through the approximations of ab initio multiple spawning. *J. Chem. Phys.* 148 (13): 134110.

45 Mignolet, B. and Curchod, B.F.E. (2019). Excited-state molecular dynamics triggered by light pulses – ab initio multiple spawning versus trajectory surface hopping. *J. Phys. Chem. A* 123: 3582.

46 Mignolet, B. and Curchod, B.F.E. (2019). Steering the outcome of a photochemical reaction – an in silico experiment on the H_2CSO sulfine using few-femtosecond dump pulses. *J. Chem. Phys.* 150 (10): 101101.

47 Mignolet, B., Curchod, B.F.E., and Martínez, T.J. (2016). Communication: XFAIMS–external field ab initio multiple spawning for electron-nuclear dynamics triggered by short laser pulses. *J. Chem. Phys.* 145 (19): 191104.

48 Mignolet, B., Curchod, B.F.E., and Martínez, T.J. (2016). Rich athermal ground-state chemistry triggered by dynamics through a conical intersection. *Angew. Chem. Int. Ed.* 128 (48): 15217–15220.

49 Mignolet, B., Curchod, B.F.E., Remacle, F., and Martínez, T.J. (2019). Sub-femtosecond stark control of molecular photoexcitation with near single-cycle pulses. *J. Phys. Chem. Lett.* 10: 742.

50 Min, S.K., Agostini, F., and Gross, E.K.U. (2015). Coupled-trajectory quantum-classical approach to electronic decoherence in non-adiabatic processes. *Phys. Rev. Lett.* 115 (7): 073001.

51 Min, S.K., Agostini, F., Tavernelli, I., and Gross, E.K.U. (2017). Ab initio non-adiabatic dynamics with coupled trajectories: a rigorous approach to quantum (de)coherence. *J. Phys. Chem. Lett.* 8: 3048.

52 Mori, T., Glover, W.J., Schuurman, M.S., and Martinez, T.J. (2012). Role of Rydberg states in the photochemical dynamics of ethylene. *J. Phys. Chem. A* 116 (11): 2808–2818.

53 Neville, S.P., Chergui, M., Stolow, A., and Schuurman, M.S. (2018). Ultrafast X-ray spectroscopy of conical intersections. *Phys. Rev. Lett.* 120 (24): 243001.

54 Penfold, T.J., Gindensperger, E., Daniel, C., and Marian, C.M. (2018). Spin-vibronic mechanism for intersystem crossing. *Chem. Rev.* 118 (15): 6975–7025.

55 Persico, M. and Granucci, G. (2014). An overview of non-adiabatic dynamics simulations methods, with focus on the direct approach versus the fitting of potential energy surfaces. *Theor. Chem. Acc.* 133 (9): 1–28.

56 Pijeau, S., Foster, D., and Hohenstein, E.G. (2017). Excited-state dynamics of 2-(2-hydroxyphenyl) benzothiazole: ultrafast proton transfer and internal conversion. *J. Phys. Chem. A* 121: 4595.

57 Richings, G., Polyak, I., Spinlove, K. et al. (2015). Quantum dynamics simulations using Gaussian wave packets: the MCG method. *Int. Rev. Phys. Chem.* 34 (2): 269–308.

58 Saita, K. and Shalashilin, D.V. (2012). On-the-fly ab initio molecular dynamics with multiconfigurational Ehrenfest method. *J. Chem. Phys.* 137 (22): 22A506.

59 Schreiner, P.R., Reisenauer, H.P., Romanski, J., and Mloston, G. (2010). Oxathiirane. *J. Am. Chem. Soc.* 132 (21): 7240–7241.

60 Shalashilin, D. (2009). Quantum mechanics with the basis set guided by Ehrenfest trajectories: theory and application to spin-boson model. *J. Chem. Phys.* 130: 244101.

61 Shalashilin, D.V. (2010). Nonadiabatic dynamics with the help of multiconfigurational Ehrenfest method: improved theory and fully quantum 24D simulation of pyrazine. *J. Chem. Phys.* 132 (24): 244111.

62 Shenvi, N., Subotnik, J.E., and Yang, W. (2011). Phase-corrected surface hopping: correcting the phase evolution of the electronic wave function. *J. Chem. Phys.* 135: 024101.

63 Shenvi, N., Subotnik, J.E., and Yang, W. (2011). Simultaneous-trajectory surface hopping: a parameter-free algorithm for implementing decoherence in non-adiabatic dynamics. *J. Chem. Phys.* 134: 144102.

64 Shenvi, N. and Yang, W. (2012). Achieving partial decoherence in surface hopping through phase correction. *J. Chem. Phys.* 137: 22A528.

65 Smith, B.D., Spears, K.G., and Sension, R.J. (2016). Probing the biexponential dynamics of ring-opening in 7-dehydrocholesterol. *J. Phys. Chem. A* 120 (33): 6575–6581.

66 Snyder, J.W. Jr., Curchod, B.F.E., and Martínez, T.J. (2016). GPU-accelerated state-averaged complete active space self-consistent field interfaced with ab initio multiple spawning unravels the photodynamics of provitamin d3. *J. Phys. Chem. Lett.* 7 (13): 2444–2449.

67 Snyder, J.W. Jr., Fales, B.S., Hohenstein, E.G. et al. (2017). A direct-compatible formulation of the coupled perturbed complete active space self-consistent field equations on graphical processing units. *J. Chem. Phys.* 146 (17): 174113.

68 Snyder, J.W. Jr., Hohenstein, E.G., Luehr, N., and Martínez, T.J. (2015). An atomic orbital-based formulation of analytical gradients and non-adiabatic coupling vector elements for the state-averaged complete active space self-consistent field method on graphical processing units. *J. Chem. Phys.* 143 (15): 154107.

69 Subotnik, J.E., Ouyang, W., and Landry, B.R. (2013). Can we derive Tully's surface-hopping algorithm from the semiclassical quantum Liouville equation? Almost, but only with decoherence. *J. Chem. Phys.* 139: 214107.

70 Subotnik, J.E. and Shenvi, N. (2011). Decoherence and surface hopping: when can averaging over initial conditions help capture the effects of wave packet separation? *J. Chem. Phys.* 134: 244114.

71 Subotnik, J.E. and Shenvi, N. (2011). A new approach to decoherence and momentum rescaling in the surface hopping algorithm. *J. Chem. Phys.* 134: 024105.

72 Hernández, H., Martínez, T.J., and Vaníček, J. (2013). Relation of exact Gaussian basis methods to the dephasing representation: theory and application to time-resolved electronic spectra. *J. Chem. Phys.* 139 (3): 034112.

73 Tao, H., Allison, T., Wright, T. et al. (2011). Ultrafast internal conversion in ethylene. I. The excited state lifetime. *J. Chem. Phys.* 134 (24): 244306.

74 Tao, H., Levine, B.G., and Martínez, T.J. (2009). Ab initio multiple spawning dynamics using multi-state second-order perturbation theory. *J. Chem. Phys. A* 113 (49): 13656–13662.

75 Tully, J.C. (1990). Molecular dynamics with electronic transitions. *J. Chem. Phys.* 93: 1061–1071.

76 Ufimtsev, I.S. and Martinez, T.J. (2008). Graphical processing units for quantum chemistry. *Comput. Sci. Eng.* 10 (6): 26–34.

77 Ufimtsev, I.S. and Martinez, T.J. (2008). Quantum chemistry on graphical processing units. 1. Strategies for two-electron integral evaluation. *J. Chem. Theory Comput.* 4 (2): 222–231.

78 Ufimtsev, I.S. and Martinez, T.J. (2009). Quantum chemistry on graphical processing units. 2. Direct self-consistent-field implementation. *J. Chem. Theory Comput.* 5 (4): 1004–1015.

79 Ufimtsev, I.S. and Martinez, T.J. (2009). Quantum chemistry on graphical processing units. 3. Analytical energy gradients, geometry optimization, and first principles molecular dynamics. *J. Chem. Theory Comput.* 5 (10): 2619–2628.

80 Virshup, A.M., Punwong, C., Pogorelov, T.V. et al. (2008). Photodynamics in complex environments: ab initio multiple spawning quantum mechanical/molecular mechanical dynamics. *J. Phys. Chem. B* 113 (11): 3280–3291.

81 Werner, H.-J., Knowles, P.J., Knizia, G. et al. (2012). Molpro, version 2012.1, a package of ab initio programs.

82 Worth, G.A., Robb, M., and Burghardt, I. (2004). A novel algorithm for non-adiabatic direct dynamics using variational Gaussian wave packets. *Faraday Discuss.* 127: 307–323.

83 Worth, G.A. and Cederbaum, L.S. (2004). Beyond Born-Oppenheimer: molecular dynamics through a conical intersection. *Annu. Rev. Phys. Chem.* 55: 127–158.

84 Worth, G.A., Robb, M.A., and Lasorne, B. (2008). Solving the time-dependent Schrödinger equation for nuclear motion in one step: direct dynamics of non-adiabatic systems. *Mol. Phys.* 106 (16–18): 2077–2091.

85 Yang, J., Zhu, X., Wolf, T.J. et al. (2018). Imaging CF_3I conical intersection and photodissociation dynamics with ultrafast electron diffraction. *Science* 361 (6397): 64–67.

86 Yang, S., Coe, J.D., Kaduk, B., and Martínez, T.J. (2009). An "optimal" spawning algorithm for adaptive basis set expansion in non-adiabatic dynamics. *J. Chem. Phys.* 130 (13): 04B606.

15

Ehrenfest Methods for Electron and Nuclear Dynamics

Adam Kirrander[1] *and Morgane Vacher*[2]

[1] *EaStCHEM, School of Chemistry, University of Edinburgh, David Brewster Road, EH9 3FJ, Edinburgh, United Kingdom*
[2] *Department of Chemistry - Ångström, The Theoretical Chemistry Programme, Uppsala University, Box 538, 751 21 Uppsala, Sweden*

Abstract

Rapid advances in ultrafast experiments have contributed to the rise of trajectory-based methods for describing nonadiabatic dynamics of molecular systems. In the present Chapter, we review theoretical methods that represent the time-dependent nuclear wave packet with trajectories that evolve classically on mean-field potential energy surfaces. The simplest form is the Ehrenfest method, which employs an ensemble of independent trajectories. We also present the Multi-Configurational Ehrenfest method in which Ehrenfest trajectories are used as a basis to solve the time-dependent Schrödinger equation. Strategies for overcoming the limitations of the methods are also discussed. Finally, we present applications of these simulation methods to electron dynamics induced by attosecond photoionisation and to nuclear dynamics probed by ultrafast x-ray scattering.

15.1 Introduction

The initial and critical steps in photo-chemical reactions occur on short time scales, and quantum effects such as nuclear motion on multiple electronic states, non-adiabatic couplings and conical intersections, tunneling, coherence and interference play an important role [26, 93, 118, 126]. In time-resolved pump-probe experiments, the initially excited wave packet is rather localized and the evolution of the system is probed over a comparatively short stretch of time. Femtosecond (1 fs = 10^{-15} s) lasers have allowed the experimental study of fundamental intramolecular dynamics since the late 1980s [125], by using light pulses that are shorter than the characteristic time for nuclear motion. In 2001, the "femtosecond barrier" was broken with the first synthesis of attosecond (1 as = 10^{-18} s) pulses [36, 74]. The developments in attoscience make it possible to observe electronic motion on its intrinsic timescale. Despite astonishing advances in ultrafast spectroscopies and imaging techniques [19, 38, 40, 81], the observables in most experiments only provide a partial picture of the dynamics, making simulations a critical tool for the interpretation of experiments.

The time evolution of a (non-relativistic) molecular system is determined by the time-dependent Schrödinger equation:

$$\mathrm{i}\hbar\frac{\partial}{\partial t}|\Phi(\vec{r},\vec{R},t)\rangle = \hat{H}|\Phi(\vec{r},\vec{R},t)\rangle, \tag{15.1}$$

Quantum Chemistry and Dynamics of Excited States: Methods and Applications,
First Edition. Edited by Leticia González and Roland Lindh.

where $|\Phi(\vec{r},\vec{R},t)\rangle$ is the total molecular wave function, $\vec{r}$ and $\vec{R}$ are the electronic and nuclear coordinates respectively. The problem will be solved in the framework of the Born–Oppenheimer approximation, as discussed in Part I. In brief, the Hamiltonian operator reads: $\hat{H} = \hat{T}_N + \hat{T}_e + \hat{V}$, with $\hat{T}_N = -\sum_I \frac{\hbar^2}{2M_I}\nabla_I^2$ the kinetic energy operator of the nuclei, $\hat{T}_e = -\sum_i \frac{\hbar^2}{2m_e}\nabla_i^2$ the kinetic energy operator of the electrons, and $\hat{V}$ the operator for inter-particle interactions (electron–electron, nucleus–nucleus and electron–nucleus). The latter two operators are normally combined to form the electronic Hamiltonian for fixed nuclei $\vec{R}$, $\hat{H}_e = \hat{T}_e + \hat{V}$. The indices i and I refer to the electrons and the nuclei respectively; m_e is used to denote the mass of an electron and M_I is used to denote the mass of the nucleus I.[1] Unfortunately, solving this equation and simulating quantum dynamics in molecules remains a fundamentally difficult problem, due to the non-local nature of quantum mechanics and the associated exponential scaling of the computational effort with the number of degrees of freedom [85]. The field of theoretical chemistry is thus dominated by developments of numerical and approximate methods.

Trajectory-based methods partially circumvent the exponential scaling problem by focusing efforts on the most relevant part of the phase-space [34]. This includes surface-hopping (see Chapter 16), simple Ehrenfest (as discussed in the current chapter), and methods that originate in the frozen Gaussian expansion of the nuclear wave function [34, 37, 86]. Modern implementations of the latter include the ab-initio multi-configurational Ehrenfest (AI-MCE) method [61] (this chapter), the ab-initio multiple spawning (AIMS) [54] (see Chapter 14), and the variational multi-configurational Gaussian (vMCG) [77] (see Chapter 13). The recently devised ab-initio multiple cloning (AIMC) method [57] can be seen as a hybrid of the AI-MCE and AIMS methods. A major feature that differentiates the various methods is the treatment of the nuclear motion. Some methods, like the surface hopping or AIMS methods, use a different set of basis trajectories for each electronic state, i.e., each basis trajectory evolves on a single potential energy surface at any given time. In technical terms, a "multi-set" formalism is used. In other methods, like the Ehrenfest or AI-MCE methods, a single set of basis trajectories is used to treat the dynamics in all electronic states and basis trajectories follow the gradient of a superposition of electronic states and therefore evolve on an *effective* potential energy surface: a "single-set" formalism is used.

The focus in this chapter is on the latter class of methods, more specifically on the Ehrenfest and AI-MCE methods [61, 80, 84, 85]. In sections 15.2 and 15.3, the equations of motion for the former and the latter are derived, respectively, highlighting the similarities and differences between the two. The underlying approximations are explicitly stated and the implications discussed. In section 15.4, illustrative applications of these methods to simulate electron and nuclear dynamics are presented.

15.2 Theory of the (Simple) Ehrenfest Method

The Ehrenfest method has been extensively discussed in the literature [7, 22, 23, 37, 55, 98, 106, 123]. The presented derivation follows Refs. [108, 113] closely.

1 The use of Cartesian coordinates is not necessary and other representations of internal coordinates are also possible.

15.2.1 Wave Function Ansatz

In order to derive mixed quantum-classical dynamics, the nuclear and electronic variables have to be separated. The simplest possible form is a product ansatz:

$$|\Phi(\vec{r},\vec{R},t)\rangle = |\Psi(\vec{r},t)\rangle|\Xi(\vec{R},t)\rangle. \tag{15.2}$$

The first approximation made in the Ehrenfest method is thus the factorization of the total wave function into a product of electronic and nuclear parts. It is noted that, following the work of Hunter [39], there have been theoretical derivations of an exact factorization of the total molecular wave function into a product of an electronic wave function and a nuclear wave function. There, the electronic wave function is not an eigenfunction (nor a superposition of eigenfunctions) of the traditional electronic Hamiltonian $\hat{H}_e$ but that of a somewhat more involved electronic Hamiltonian which depends on the nuclear wave function itself [5, 17]. Here, the ansatz (15.2) is called a single-configuration ansatz for the total wave function. It is emphasized that "single-configuration" refers to the total molecular wave function ansatz, and not to the electronic structure method (the latter is not discussed in this chapter).

One may expand the electronic wave function in a basis of N_s orthonormal electronic states $\{\Psi_\beta\}$ that depend parametrically on $\vec{R}$. This will be used to obtain practical equations of motion. The total wave function ansatz in Eq. (15.2) thus reads:

$$|\Phi(\vec{r},\vec{R},t)\rangle = \left(\sum_{\beta=1}^{N_s} a_\beta(t)|\Psi_\beta(\vec{r};\vec{R})\rangle\right)|\Xi(\vec{R},t)\rangle. \tag{15.3}$$

Each of the electronic eigenfunctions is multiplied by a complex amplitude, $a_\beta(t)$, such that $|a_\beta(t)|^2$ is the population on the corresponding electronic state β. Since the electronic basis wave functions are orthonormal $\langle\Psi_\beta(\vec{r}\,;\vec{R})|\Psi_{\beta'}(\vec{r}\,;\vec{R})\rangle_{\vec{r}} = \delta_{\beta\beta'}$ and $1 = \sum_{\beta=1}^{N_s}|a_\beta(t)|^2$, the total wave function is normalized (assuming a normalized nuclear wave function $\langle\Xi(\vec{R},t)|\Xi(\vec{R}),t\rangle_{\vec{R}} = 1$). The subscripts $\vec{r}$ and $\vec{R}$ indicate the coordinate of integration.

It is noted that the single-configuration ansatz (15.2) is different from the Born–Oppenheimer ansatz [10] for separating the electronic and nuclear variables even in its one-determinant limit, where only a single electronic eigenstate of $\hat{H}_e$ would be included in the expansion. Here, the electronic wave function $|\Psi(\vec{r},t)\rangle$ does not have to be a single adiabatic state.

One deficiency of the ansatz (15.2) is the fact that the electronic wave function does not have the possibility to decohere: the populated electronic states in $|\Psi(\vec{r},t)\rangle$ share the same nuclear wave packet $|\Xi(\vec{R},t)\rangle$ by definition of the total wave function. Decoherence here is defined as the tendency of the time-evolved electronic wave function to behave as a statistical ensemble of electronic states rather than a coherent superposition of them [127]. The neglect of electronic decoherence could lead to non-physical asymptotic behavior in the case of bifurcating paths, for instance into different product channels during a chemical reaction.

In order to simplify the appearance of the expressions at a later stage of the derivation [106], a phase factor is introduced for the total wave function in Eq. (15.2),

$$|\Phi(\vec{r},\vec{R},t)\rangle = |\Psi(\vec{r},t)\rangle|\Xi(\vec{R},t)\rangle\exp\left(\frac{\mathrm{i}}{\hbar}\int^t E(t')dt'\right), \tag{15.4}$$

and also some internal phase factors for the two individual wave functions,

$$\mathrm{i}\hbar\left\langle\Xi\left|\frac{\partial\Xi}{\partial t}\right.\right\rangle_{\vec{R}} = E^{tot} \text{ and } \mathrm{i}\hbar\left\langle\Psi\left|\frac{\partial\Psi}{\partial t}\right.\right\rangle_{\vec{r}} = E(t), \tag{15.5}$$

with $E(t) = \langle\Xi\Psi|\hat{H}_e|\Xi\Psi\rangle_{\vec{R},\vec{r}}$ and $E^{tot} = \langle\Xi\Psi|\hat{H}|\Xi\Psi\rangle_{\vec{R},\vec{r}}$.

15.2.2 Equations of Motion

Inserting the ansatz (15.4) with the additional phase factors into Eq. (15.1), and closing on the left by $\langle\Xi(\vec{R},t)|$ and $\langle\Psi(\vec{r},t)|$ gives respectively:

$$\mathrm{i}\hbar\frac{\partial|\Psi(\vec{r},t)\rangle}{\partial t} = \hat{T}_e|\Psi(\vec{r},t)\rangle + \langle\Xi(\vec{R},t)|\hat{V}(\vec{r},\vec{R})|\Xi(\vec{R},t)\rangle_{\vec{R}}|\Psi(\vec{r},t)\rangle \tag{15.6}$$

$$\mathrm{i}\hbar\frac{\partial|\Xi(\vec{R},t)\rangle}{\partial t} = \hat{T}_N|\Xi(\vec{R},t)\rangle + \langle\Psi(\vec{r},t)|\hat{H}_e(\vec{r};\vec{R})|\Psi(\vec{r},t)\rangle_{\vec{r}}|\Xi(\vec{R},t)\rangle. \tag{15.7}$$

The set of coupled Eqs. (15.6) and (15.7) are the basis of the time-dependent self-consistent field (TD-SCF) method [24, 32], also referred to as time-dependent Hartree (TDH) when the nuclear wave function $|\Xi(\vec{R},t)\rangle$ is written as a simple product of (time-dependent) one-dimensional functions. The same equations could have been obtained by applying the Dirac–Frenkel variational principle [113]. The mean-field origin of the TD-SCF approach imposes limitations, as already discussed above. By looking at the second term on the right-hand side of Eq. (15.6), we see that the interaction between electrons at points $\vec{r}$ and nuclei at points $\vec{R}$ is weighted by the probability that the nuclei are at these particular points. This is the *effective* potential experienced by the electrons due to the nuclei. The corresponding comment can be made about the second term on the right-hand side of Eq. (15.7). According to the set of coupled equations (15.6) and (15.7), both electrons and nuclei move thus in time-dependent effective potentials obtained from appropriate expectation values of the nuclear and electronic wave functions respectively. In other words, the feedback between electronic and nuclear degrees of freedom is described in a mean-field manner, in both directions.

The Ehrenfest method is the classical analogue to the TD-SCF method [28] and therefore inherits the same limitation. It is obtained by taking the classical limit for nuclear motion of Eqs. (15.6) and (15.7). To do that in Eq. (15.7), the nuclear wave function is (exactly) rewritten as,

$$|\Xi(\vec{R},t)\rangle = \exp\left(\frac{\mathrm{i}}{\hbar}S(\vec{R},t)\right), \tag{15.8}$$

allowing for a complex phase S [100]. After inserting ansatz (15.8) in (15.7), we obtain:

$$\frac{\partial S}{\partial t} + \sum_I \frac{1}{2M_I}(\vec{\nabla}_I S)^2 + \langle\Psi(\vec{r},t)|\hat{H}_e(\vec{r};\vec{R})|\Psi(\vec{r},t)\rangle_{\vec{R}} = \mathrm{i}\hbar\sum_I \frac{1}{2M_I}\nabla_I^2 S. \tag{15.9}$$

Equation (15.9) is equivalent to the original Eq. (15.7). The right-hand side term (proportional to $\hbar$) may be thought of as a time-dependent "quantum correction". The classical Hamilton–Jacobi equation [30], one of many possible formulations of classical mechanics, is obtained when taking the limit $\hbar \to 0$:

$$\frac{\partial S}{\partial t} + \sum_I \frac{1}{2M_I}(\vec{\nabla}_I S)^2 + \langle\Psi(\vec{r},t)|\hat{H}_e(\vec{r};\vec{R})|\Psi(\vec{r},t)\rangle_{\vec{r}} = 0. \tag{15.10}$$

The resulting Eq. (15.10) is thus equivalent to Newton's equation of motion, where $\vec{P} = \vec{\nabla}_I S = \frac{\partial S}{\partial \vec{R}}$ gathers the classical nuclear momentum coordinates:

$$\frac{\mathrm{d}\vec{P}}{\mathrm{d}t} = \vec{F} = -\frac{\partial}{\partial\vec{R}}\langle\Psi(\vec{r},t)|\hat{H}_e(\vec{r};\vec{R})|\Psi(\vec{r},t)\rangle_{\vec{r}}, \tag{15.11}$$

where $\vec{F}$ is the mean-field force experienced by the nuclei due to the electronic wave packet $|\Psi(\vec{r},t)\rangle$.

Expanding the electronic wave function in terms of the basis of electronic states as in Eq. (15.3), the expression for the force (15.11) becomes:

$$\vec{F} = -\sum_{\beta,\beta'} a^*_\beta(t)a_{\beta'}(t)\frac{\partial}{\partial\vec{R}}\langle\Psi_\beta(\vec{r};\vec{R})|\hat{H}_e(\vec{r};\vec{R})|\Psi_{\beta'}(\vec{r};\vec{R})\rangle_{\vec{r}}. \tag{15.12}$$

In the case of a basis of adiabatic states (as provided by electronic structure packages), it reads:

$$\vec{F} = -\sum_{\beta} |a_\beta(t)|^2 \frac{\mathrm{d}E_\beta(\vec{R})}{\mathrm{d}\vec{R}} - \sum_{\beta \neq \beta'} a^*_\beta(t) a_{\beta'}(t) \vec{d}_{\beta\beta'}(\vec{R})(E_{\beta'}(\vec{R}) - E_\beta(\vec{R})), \tag{15.13}$$

with $E_\beta(\vec{R})$ the electronic eigenvalues and $\vec{d}_{\beta\beta'}(\vec{R}) = \langle \Psi_\beta | \frac{\partial}{\partial \vec{R}} \Psi_{\beta'} \rangle = \frac{\langle \Psi_\beta | (\frac{\partial}{\partial \vec{R}} \hat{H}_e) | \Psi_{\beta'} \rangle}{E_{\beta'}(\vec{R}) - E_\beta(\vec{R})}$ the non-adiabatic coupling matrix elements (NACMEs). The first sum on the right-hand side of Eq. (15.13) gives the nuclear gradients of the individual electronic basis states weighted by their population. The second sum accounts for the change in population because of non-adiabatic transitions; it determines the component of the gradient in the direction of the derivative couplings. This means that the effective force felt by the nuclei contains contributions from the non-adiabatic couplings.

To integrate the equation of motion for the nuclear geometry, one could also directly compute the nuclear gradient of the global electronic wave packet $|\Psi(\vec{r}, t)\rangle$ and use Eq. (15.11), instead of computing the gradient for each individual electronic state as well as the NACMEs for each pair of states. This is one of the advantages of the Ehrenfest approach. A general approach to compute the gradient and Hessian of configuration interaction-like wave functions is the method of Almlöf and Taylor [6]. For more detailed information regarding this, the reader is referred to Refs. [44, 108]. Here, as in other methods, one can make a local harmonic approximation of the *effective* potential energy surface around a point $\vec{R}_0$ using the first and second derivatives of the energy of the electronic wave packet:

$$E(\vec{R}) = E^0 + \vec{G}^0 \cdot (\vec{R} - \vec{R}_0) + \frac{1}{2}(\vec{R} - \vec{R}_0)^\dagger \cdot \mathbf{H}^0 \cdot (\vec{R} - \vec{R}_0), \tag{15.14}$$

where $E^0 = E(\vec{R}_0)$, $\vec{G}^0 = \left.\frac{\mathrm{d}E(\vec{R})}{\mathrm{d}\vec{R}}\right|_{\vec{R}_0}$ and $\mathbf{H}^0 = \left.\frac{\mathrm{d}^2E(\vec{R})}{\mathrm{d}\vec{R}^2}\right|_{\vec{R}_0}$ are the energy, the gradient and the Hessian evaluated at $\vec{R}^0$, respectively. Newton's equation of motion on a quadratic surface is:

$$\frac{\mathrm{d}\vec{P}}{\mathrm{d}t} = -\vec{G}^0 - \mathbf{H}^0 \cdot (\vec{R} - \vec{R}_0). \tag{15.15}$$

The velocity Verlet algorithm or any other propagator may then be used to integrate Eq. (15.15). However, if a Hessian is included, a higher order propagator should be used.

In the electronic equation of motion (15.6), the classical limit is taken by replacing $|\Xi(\vec{R}, t)\rangle$ by a delta function at the classical trajectory $\vec{R}(t)$:

$$\begin{aligned} \mathrm{i}\hbar \frac{\partial |\Psi(\vec{r}, t; \vec{R})\rangle}{\partial t} &= \left(-\sum_i \frac{\hbar^2}{2m_e} \frac{\partial^2}{\partial \vec{r}^2} + \hat{V}(\vec{r}, \vec{R}(t)) \right) |\Psi(\vec{r}, t; \vec{R})\rangle \\ &= \hat{H}_e(\vec{r}; \vec{R}(t)) |\Psi(\vec{r}, t; \vec{R})\rangle. \end{aligned} \tag{15.16}$$

Here, the electronic wave function depends parametrically on $\vec{R}(t)$ through $\hat{V}(\vec{r}, \vec{R}(t))$ and thus $\hat{H}_e(\vec{r}; \vec{R}(t))$.

Substituting the electronic basis expansion (15.3) into Eq. (15.16) and closing on the left by $\langle \Psi_\alpha(\vec{r}; \vec{R})|$ gives:

$$\frac{\mathrm{d}a_\alpha(t)}{\mathrm{d}t} = -\frac{\mathrm{i}}{\hbar} \sum_\beta a_\beta(t) \langle \Psi_\alpha | \hat{H}_e | \Psi_\beta \rangle - \sum_\beta a_\beta(t) \left\langle \Psi_\alpha \middle| \frac{\partial}{\partial t} \Psi_\beta \right\rangle. \tag{15.17}$$

Eq. (15.17) gives the time-dependent amplitude coefficients of the electronic states. Note that the same equations define the time-dependence of the electronic basis amplitudes for the surface hopping method.

One can work in the adiabatic representation using the electronic eigenstates. Then, the sum in the first term on the right-hand side of Eq. (15.17) disappears since the electronic Hamiltonian matrix is diagonal in the adiabatic basis:

$$\frac{\mathrm{d}a_\alpha(t)}{\mathrm{d}t} = -\frac{\mathrm{i}}{\hbar} a_\alpha(t) E_\alpha(\vec{R}) - \sum_{\beta\neq\alpha} a_\beta(t) \vec{d}_{\alpha\beta}(\vec{R}) \cdot \frac{\mathrm{d}\vec{R}}{\mathrm{d}t}. \tag{15.18}$$

The NACMEs present a singularity at the point of degeneracy of the electronic eigenstates, which may cause numerical problems. One can avoid the singularity problem by working in a different representation where the electronic Hamiltonian is not necessarily diagonal but where the non-adiabatic coupling is (approximately) zero. Such a basis is called a (*quasi-*)*diabatic* basis [21, 56, 65]. Equation (15.17) then reads:

$$\frac{\mathrm{d}a_\alpha(t)}{\mathrm{d}t} = -\frac{\mathrm{i}}{\hbar} \sum_\beta a_\beta(t) \langle \Psi_\alpha | \hat{H}_e | \Psi_\beta \rangle. \tag{15.19}$$

The coupling between electronic states is now due to the off-diagonal elements of the electronic Hamiltonian matrix. Equations (15.11) and (15.16), and equivalently (15.13) and (15.18) using the electronic adiabatic basis set, define the Ehrenfest method. One advantage of the Ehrenfest method is that it is, in principle, representation-independent; its applications and results do not depend on the choice of electronic basis functions (if complete). By treating the nuclear motion classically, we lose the spatial delocalization of the nuclear wave packet: the nuclear motion is now described by a classical trajectory, i.e., by a line in phase space. To obtain a realistic description of the dynamics of the system, one mimics the nuclear wave packet motion by propagating a swarm of independent classical trajectories starting with sampled positions $\vec{R}$ and momenta $\vec{P}$ of the nuclei. Because of the mean-field treatment of the interactions between electrons and nuclei, the Ehrenfest method succeeds in describing nuclear motion if the potential energy surfaces of the various electronic states are similar in topology and energies [41]. However, in the case of weakly coupled electronic states, the nuclear motion will be dominated by the potential corresponding to the highly populated electronic state and regions of space accessible only on the sparsely populated electronic state may not be explored properly [31, 105]. Non-physical asymptotic behaviors could be obtained in the case of bifurcating paths. A practical approach to overcome this problem will be presented below.

15.3 Theory of the Multi-Configurational Ehrenfest Method

In this section, we review the multi-configurational Ehrenfest (MCE) method and how it builds on the (simple) Ehrenfest method presented above. The variants considered here include *Ab-Initio* MCE (AI-MCE) [80, 84, 85] and *Ab-Initio* Multiple Cloning MCE (AIMC-MCE) [57, 58]. These are closely related to the coupled coherent states (CCS) method [86] and *Ab-Initio* Multiple Spawning (AIMS) method [62, 64].

15.3.1 Wave Function Ansatz

In the multi-configurational Ehrenfest method, the total molecular wave function $|\Phi(\vec{r},\vec{R},t)\rangle$ is expanded as a sum of N_{trj} Ehrenfest wave packets $|\Phi_\mu(\vec{r},\vec{R},t)\rangle$ with complex coefficients $D_\mu(t)$ [84]:

$$|\Phi(\vec{r},\vec{R},t)\rangle = \sum_{k=1}^{N_{\mathrm{trj}}} D_\mu(t) |\Phi_\mu(\vec{r},\vec{R},t)\rangle. \tag{15.20}$$

Each Ehrenfest wave packet reads,

$$|\Phi_\mu(\vec{r},\vec{R},t)\rangle = \left[\sum_{\beta=1}^{N_s} a_\mu^\beta(t)|\Psi_\beta(\vec{r}\,;\vec{R})\rangle\right]|g_\mu(\vec{R},t)\rangle, \tag{15.21}$$

where N_s is the number of electronic states $|\Psi_\beta(\vec{r}\,;\vec{R})\rangle$ considered. Equation (15.21) is equivalent to the ansatz (15.3), where the nuclear wave packet $|\Xi(\vec{R},t)\rangle$ is now a normalized frozen Gaussian wave packet $|g_\mu(\vec{R},t)\rangle$, akin to a constant width Gaussian in a harmonic potential. This is also known as a coherent state, and given by,

$$|g_\mu(\vec{R},t)\rangle = \left(\frac{\alpha}{\pi}\right)^{N_{\text{dof}}/4} \exp\left(-\frac{\alpha}{2}[\vec{R}-\vec{R}_\mu(t)]^2 + \frac{\mathrm{i}}{\hbar}\vec{P}_\mu(t)[\vec{R}-\vec{R}_\mu(t)] + \frac{\mathrm{i}}{\hbar}\frac{\vec{P}_\mu(t)\vec{R}_\mu(t)}{2}\right), \tag{15.22}$$

where N_{dof} is the number of nuclear degrees of freedom, $\vec{R}_\mu(t)$ the position and $\vec{P}_\mu(t)$ the momentum coordinates of the wave packet in phase-space, and the final term is the semiclassical phase γ_μ [86],

$$\gamma_\mu(t) = \frac{\vec{P}_\mu(t)\vec{R}_\mu(t)}{2}. \tag{15.23}$$

The width-parameter α is fixed, according to the frozen Gaussian approximation [33], and can be specified individually for each degree of freedom, for instance according to the type of atom and chemical context using a prescription elaborated in Ref. [103]. Further context on the width parameters provided is in Ref. [49]. The nuclear probability distribution of the different coherent states is then given by,

$$|g_\mu(\vec{R},t)|^2 = \left(\frac{\alpha}{\pi}\right)^{N_{\text{dof}}/2} \exp(-\alpha[\vec{R}-\vec{R}_\mu(t)]^2), \tag{15.24}$$

which is a Gaussian-shaped distribution around the central coordinate $\vec{R}_\mu(t)$. Each coherent state is normalized, $\langle g_\mu(\vec{R},t)|g_\mu(\vec{R},t)\rangle = 1$, and the overlap of two coherent states is,

$$\Omega_{\mu'\mu}(t) = \langle g_{\mu'}(\vec{R},t)|g_\mu(\vec{R},t)\rangle = \exp\left(\vec{Z}^*_{\mu'}(t)\vec{Z}_\mu(t) - \frac{\vec{Z}^*_{\mu'}(t)\vec{Z}_{\mu'}(t)}{2} - \frac{\vec{Z}^*_\mu(t)\vec{Z}_\mu(t)}{2}\right), \tag{15.25}$$

where $\vec{Z}_\mu(t) = \alpha^{1/2}\vec{R}_\mu(t) + \mathrm{i}\alpha^{-1/2}\hbar^{-1}\vec{P}_\mu(t)$ (and analogously for $\vec{Z}_{\mu'}(t)$) is a compound coordinate that allows for compact representation of phase-space coordinates [86]. The overlap matrix for the Ehrenfest wave packets defined in Eq. (15.21) is similar, but includes the electronic components,

$$\tilde{\Omega}_{\mu'\mu}(t) = \langle\Phi_{\mu'}(\vec{r},\vec{R},t)|\Phi_\mu(\vec{r},\vec{R},t)\rangle = \left(\sum_\beta a_{\mu'}^{\beta*}(t)a_\mu^\beta(t)\right)\Omega_{\mu'\mu}(t). \tag{15.26}$$

As a consequence of Eqs. (15.25) and (15.26) it is clear that the Ehrenfest states are not orthonormal. The identity operator $\hat{1}$ must account for this non-orthonormality and is given by,

$$\hat{1} = \sum_{\beta\gamma}|\Phi_\beta(\vec{r},\vec{R},t)\rangle\,\tilde{\Omega}^{-1}_{\beta\gamma}(t)\,\langle\Phi_\gamma(\vec{r},\vec{R},t)|. \tag{15.27}$$

The norm of the total molecular wave function in Eq. (15.20) is therefore calculated as $\langle\Phi_\mu(\vec{r},\vec{R},t)|\hat{1}|\Phi_\mu(\vec{r},\vec{R},t)\rangle$ with $\hat{1}$ as defined in Eq. (15.27) above.

15.3.2 Equations of Motion

The time evolution of the total molecular wave function, $|\Phi(\vec{r},\vec{R},t)\rangle$ in Eq. (15.20), is given by the expansion coefficients $D_\mu(t)$ and the propagation of each individual wave packet $|\Phi_\mu(\vec{r},\vec{R},t)\rangle$. The quantum mechanically most appropriate manner of propagation is obtained using the variational principle [52] as is done in the vMCG [77, 119, 120] and G-MCTDH [13] methods. However, the resulting equations do not parallelize well since all variables (positions, momenta, and amplitudes) become coupled and solving the equations may become numerically unstable. Therefore, MCE takes a simpler approach where each individual wave packet $|\Phi_\mu(\vec{r},\vec{R},t)\rangle$ is guided along an independent semiclassical trajectory in phase-space according to classical equations for that specific wave packet only. In the following, we derive the equations of motion for the MCE approach, starting with the expansion coefficients $D_\mu(t)$, then proceeding to the amplitudes $a_\mu^\beta(t)$ and phase-space coordinates $(\vec{R}_\mu(t),\vec{P}_\mu(t))$ of the individual wave packets. The numerical solution of the equations of motion ultimately requires electronic potential energies, gradients and non-adiabatic couplings from ab-initio electronic structure calculations. For a more detailed derivation of variant forms of MCE, including for instance those formulated in a diabatic basis, the reader is directed to a review by Makhov *et al* [61] (for adiabatic versus diabatic representations see Part I). Notably, MCE appears in two main versions referred to as MCEv1 and MCEv2 [61]. The focus in this chapter is entirely on MCEv2 which was developed specifically for ab-initio direct dynamics. Here, the trajectories are propagated independently and only coupled a posteriori via the time-dependent expansion coefficients $D_\mu(t)$. In contrast, in MCEv1 all trajectories are coupled throughout the calculation, with each trajectory influenced by all other trajectories. This gives good convergence [83], but is impractical for ab-initio direct dynamics.

The time-evolution of the coefficients $D_\mu(t)$ is due to a posteriori coupling between different trajectories and is obtained by substituting the ansatz in Eq. (15.20) into the time-dependent Schrödinger Eq. (15.1) and closing by $\langle\Phi_{\mu'}(\vec{r},\vec{R},t)|$ on the left,

$$\sum_\mu \tilde{\Omega}_{\mu'\mu}\frac{\mathrm{d}D_\mu(t)}{\mathrm{d}t} = -\frac{\mathrm{i}}{\hbar}\sum_\mu\Bigg(\langle\Phi_{\mu'}(\vec{r},\vec{R},t)|\hat{H}|\Phi_\mu(\vec{r},\vec{R},t)\rangle - \mathrm{i}\left\langle\Phi_{\mu'}(\vec{r},\vec{R},t)\middle|\frac{\partial\Phi_\mu(\vec{r},\vec{R},t)}{\partial t}\right\rangle\Bigg)D_\mu(t). \tag{15.28}$$

Solving Eq. (15.28) requires the overlap between two wave packets, which appears on the left-hand side of Eq. (15.28) and which was obtained in Eq. (15.26) previously, as well as the following two terms which appear on the right-hand side of Eq. (15.28):

$$\langle\Phi_{\mu'}(\vec{r},\vec{R},t)|\hat{H}|\Phi_\mu(\vec{r},\vec{R},t)\rangle = \sum_{\beta\beta'}\langle g_{\mu'}(\vec{R},t)|\langle\Psi_\beta(\vec{r}\,;\vec{R})|\hat{H}|\Psi_{\beta'}(\vec{r}\,;\vec{R})\rangle|g_\mu(\vec{R},t)\rangle a_{\mu'}^{\beta*}(t)a_\mu^{\beta'}(t), \tag{15.29}$$

and

$$\left\langle\Phi_{\mu'}(\vec{r},\vec{R},t)\middle|\frac{\partial\Phi_\mu(\vec{r},\vec{R},t)}{\partial t}\right\rangle = \left\langle g_{\mu'}(\vec{R},t)\middle|\frac{\partial g_\mu(\vec{R},t)}{\partial t}\right\rangle\sum_\beta a_{\mu'}^{\beta*}(t)a_\mu^\beta(t) + \left\langle g_{\mu'}(\vec{R},t)\middle|g_\mu(\vec{R},t)\right\rangle\sum_\beta a_{\mu'}^{\beta*}(t)\frac{\mathrm{d}a_\mu^\beta(t)}{\mathrm{d}t}. \tag{15.30}$$

We proceed to find more detailed expressions for Eqs. (15.29, 15.30), beginning with the Hamiltonian matrix element on the right-hand side of Eq. (15.29). We recall that $\hat{H} = \hat{T}_N + \hat{H}_e$, and consider each of the contributions from $\hat{T}_N$ and $\hat{H}_e$ separately.

To calculate the contribution due to the electronic Hamiltonian $\hat{H}_e$ we note that the electronic potential terms are diagonal in an adiabatic basis,

$$\langle \Psi_\beta(\vec{r}\,;\vec{R})|\hat{H}_\mathrm{e}|\Psi_{\beta'}(\vec{r}\,;\vec{R})\rangle = E_\beta(\vec{R})\;\delta_{\beta\beta'}. \tag{15.31}$$

The computational cost of the diagonal potential matrix elements is minimized by approximating the potential with the value at the center of the coherent state,

$$\langle g_\mu(\vec{R},t))|E_\beta|g_\mu(\vec{R},t)\rangle \approx E_\beta(\vec{R}_\mu). \tag{15.32}$$

The corresponding off-diagonal matrix elements must be calculated using at least a first-order approximation with respect to the coherent state [61],

$$\begin{aligned}\langle g_{\mu'}(\vec{R},t))|E_\beta|g_\mu(\vec{R},t)\rangle \approx\; & \frac{1}{2}\langle g_{\mu'}(\vec{R},t)|g_\mu(\vec{R},t)\rangle[E_\beta(\vec{R}_{\mu'}) + E_\beta(\vec{R}_\mu)] \\ & + \frac{1}{2}\langle g_{\mu'}(\vec{R},t)|(\vec{R}-\vec{R}_{\mu'})|g_\mu(\vec{R},t)\rangle\frac{\mathrm{d}E_\beta(\vec{R}_{\mu'})}{\mathrm{d}\vec{R}_{\mu'}} \\ & + \frac{1}{2}\langle g_{\mu'}(\vec{R},t)|(\vec{R}-\vec{R}_{\mu})|g_\mu(\vec{R},t)\rangle\frac{\mathrm{d}E_\beta(\vec{R}_{\mu})}{\mathrm{d}\vec{R}_{\mu}}.\end{aligned} \tag{15.33}$$

This first-order approximation only uses the values of energies and gradients at the central coordinates $\vec{R}_\mu$ of the wave packets, which are always calculated during the propagation, meaning that the coupling between wave packets is obtained with little computational effort.

Subsequent evaluation of the contribution to Eq. (15.29) from the kinetic energy operator $\hat{T}_N$ yields the derivative coupling due to the parametric $\vec{R}$-dependence of the electronic wave functions responsible for non-adiabatic population transfer between electronic states. Ignoring the second-derivative terms since they only have a small impact on the dynamics [2, 47, 63] results in,

$$\begin{aligned}&\langle g_\mu(\vec{R},t)|\langle\Psi_\beta(\vec{r}\,;\vec{R})|\;\hat{T}_N\;|\Psi_{\beta'}(\vec{r}\,;\vec{R})\rangle|g_\mu(\vec{R},t)\rangle \\ &\qquad\approx\; -\frac{\hbar^2}{2\mathbf{M}}\left\langle g_\mu(\vec{R},t)\middle|\left\langle \Psi_\beta\vec{r}\,;\vec{R})|\Psi_{\beta'}\vec{r}\,;\vec{R})\right\rangle\middle|\frac{\partial^2 g_\mu(\vec{R},t)}{\partial\vec{R}^2}\right\rangle \\ &\qquad\quad -\frac{\hbar^2}{\mathbf{M}}\left\langle g_\mu(\vec{R},t)\middle|\left\langle \Psi_\beta\vec{r}\,;\vec{R})\middle|\frac{\partial\Psi_{\beta'}\vec{r}\,;\vec{R})}{\partial\vec{R}}\right\rangle\middle|\frac{\partial g_\mu}{\partial\vec{R}}\right\rangle \\ &\qquad\approx\; \frac{\vec{P}_\mu(t)^2}{2\mathbf{M}}\delta_{\beta\beta'} - \mathrm{i}\hbar\frac{\vec{P}_\mu(t)^2}{\mathbf{M}}\vec{d}_{\beta\beta'}(\vec{R}_\mu),\end{aligned} \tag{15.34}$$

where $\mathbf{M}$ is a diagonal matrix with the masses of each associated degree of freedom along the diagonal. For the matrix elements of $\hat{T}_N$ between different trajectories, a simple approximation is used,

$$\begin{aligned}&\langle g_{\mu'}(\vec{R},t)|\langle\Psi_\beta(\vec{r}\,;\vec{R})|\;\hat{T}_N\;|\Psi_{\beta'}(\vec{r}\,;\vec{R})\rangle|g_\mu(\vec{R},t)\rangle \\ &\qquad\approx\; \langle g_{\mu'}(\vec{R},t)|\hat{T}_N|g_\mu(\vec{R},t)\rangle\delta_{\beta\beta'} \\ &\qquad\quad -\mathrm{i}\hbar\frac{\vec{P}_\mu(t)}{2\mathbf{M}}\left(\vec{d}_{\beta\beta'}(\vec{R}_{\mu'}) + \vec{d}_{\beta\beta'}(\vec{R}_\mu)\right),\end{aligned} \tag{15.35}$$

where the following expression concludes our exploration of the contributions to Eq. (15.29),

$$\langle g_{\mu'}(\vec{R},t)|\hat{T}_N|g_\mu(\vec{R},t)\rangle = \frac{\langle g_{\mu'}(\vec{R},t)|g_\mu(\vec{R},t)\rangle}{2\mathbf{M}}\left[\left(\frac{\vec{P}_\mu(t)+\vec{P}_{\mu'}(t)}{2}\right)^2 + \alpha\hbar^2 - \alpha^2\hbar^2(\vec{R}_\mu(t)-\vec{R}_{\mu'}(t))^2 - \mathrm{i}\alpha\hbar(\vec{R}_\mu(t)-\vec{R}_{\mu'}(t))(\vec{P}_\mu(t)+\vec{P}_{\mu'}(t))\right]. \tag{15.36}$$

We proceed to find an expression for the derivative overlap between two coherent states that appears on the right-hand side of Eq. (15.30),

$$\left\langle g_\mu(\vec{R},t)\left|\frac{\partial g_\mu(\vec{R},t)}{\partial t}\right.\right\rangle = \frac{\mathrm{d}\vec{R}_\mu}{\mathrm{d}t}\left\langle g_\mu(\vec{R},t)\left|\frac{\partial g_\mu(\vec{R},t)}{\partial \vec{R}_\mu}\right.\right\rangle + \frac{\mathrm{d}\vec{P}_\mu}{\mathrm{d}t}\left\langle g_\mu(\vec{R},t)\left|\frac{\partial g_\mu(\vec{R},t)}{\partial \vec{P}_\mu}\right.\right\rangle + \frac{\mathrm{i}}{\hbar}\frac{\mathrm{d}\gamma_\mu}{\mathrm{d}t}, \tag{15.37}$$

where the evolution of the phase $\gamma_\mu(t)$, which was defined in Eq. (15.23) previously, is given by [86],

$$\frac{\mathrm{d}\gamma_\mu(t)}{\mathrm{d}t} = \frac{\vec{P}_\mu}{2}\frac{\mathrm{d}\vec{R}_\mu}{\mathrm{d}t}. \tag{15.38}$$

Then, using the following expressions for the matrix elements,

$$\left\langle g_\mu(\vec{R},t)\left|\frac{\partial g_\mu(\vec{R},t)}{\partial \vec{R}_\mu}\right.\right\rangle = -\frac{\mathrm{i}}{\hbar}\vec{P}_\mu(t), \tag{15.39}$$

and

$$\left\langle g_\mu(\vec{R},t)\left|\frac{\partial g_\mu(\vec{R},t)}{\partial \vec{P}_\mu}\right.\right\rangle = 0, \tag{15.40}$$

we obtain

$$\left\langle g_\mu(\vec{R},t)\left|\frac{\partial g_\mu(\vec{R},t)}{\partial t}\right.\right\rangle = -\frac{\mathrm{i}}{2\hbar}\frac{\mathrm{d}\vec{R}_\mu}{\mathrm{d}t}\vec{P}_\mu(t), \tag{15.41}$$

which provides the final component for the evaluation of Eq. (15.28). Having thus defined all components required to propagate the expansion coefficients $D_\mu(t)$ via Eq. (15.28), we proceed to the equations of motion for the Ehrenfest wave packets.

In the propagation of the individual Ehrenfest wave packets, we have to consider the amplitudes $a_\mu^\beta(t)$ and phase-space coordinates $(\vec{R}_\mu(t),\vec{P}_\mu(t))$. We begin with the amplitudes $a_\mu^\beta(t)$, which can be obtained by considering a single Ehrenfest wave packet in the time-dependent Schrödinger equation, analogously to what was done previously in Eq. (15.28) for the $D_\mu(t)$ coefficients. This leads to the following equation [61],

$$\frac{\mathrm{d}a_\mu^\beta(t)}{\mathrm{d}t} = -\frac{\mathrm{i}}{\hbar}\sum_{\beta'}\langle g_\mu(\vec{R},t)|\langle\Psi_\beta(\vec{r}\,;\vec{R})|\hat{H}|\Psi_{\beta'}(\vec{r}\,;\vec{R})\rangle|g_\mu(\vec{R},t)\rangle a_\mu^{\beta'}(t) - \left\langle g_\mu(\vec{R},t)\left|\frac{\partial g_\mu(\vec{R},t)}{\partial t}\right.\right\rangle a_\mu^\beta(t). \tag{15.42}$$

Substituting Eqs. (15.31) and (15.41) and the approximations in Eqs. (15.32) and (15.34) into Eq. (15.42), we obtain,

$$\frac{\mathrm{d}a_\mu^\beta(t)}{\mathrm{d}t} = -\frac{\mathrm{i}}{\hbar} a_\mu^\beta(t) E_\beta(\vec{R}_\mu) - \sum_{\beta' \neq \beta} a_\mu^{\beta'}(t) \vec{d}_{\beta\beta'}(\vec{R}_\mu) \cdot \frac{\mathrm{d}\vec{R}_\mu}{\mathrm{d}t}. \tag{15.43}$$

This is identical to Eq. (15.18) derived in the (simple) Ehrenfest method. The absence of coupling between the amplitudes $a_\mu^\beta(t)$ for different Ehrenfest wave packets is important, as it simplifies the computational procedures making this multi-configurational approach suitable for direct ab-initio dynamics.

Finally, we must obtain the forces used to classically propagate the phase-space coordinates $(\vec{R}_\mu(t), \vec{P}_\mu(t))$ of each Ehrenfest wave packet. Direct substitution of the Ehrenfest ansatz yields the average (mean-field) Ehrenfest force as,

$$\begin{aligned}\vec{F}_\mu &= -\sum_{\beta\beta'} a_\mu^{\beta*}(t) a_\mu^{\beta'}(t) \frac{\partial}{\partial \vec{R}_\mu} \langle g_\mu(\vec{R},t) | \langle \Psi_\beta(\vec{r}\,;\vec{R}) | \hat{H}_e | \Psi_{\beta'}(\vec{r}\,;\vec{R}) \rangle | g_\mu(\vec{R},t) \rangle \\ &= -\sum_\beta |a_\mu^\beta(t)|^2 \frac{\mathrm{d}E_\beta(\vec{R}_\mu)}{\mathrm{d}\vec{R}_\mu} - \sum_{\beta \neq \beta'} a_\mu^{\beta*}(t) a_\mu^{\beta'}(t)\, \vec{d}_{\beta\beta'}(\vec{R}_\mu)\, (E_{\beta'}(\vec{R}_\mu) - E_\beta(\vec{R}_\mu)),\end{aligned} \tag{15.44}$$

where we make use of the fact that the coherent states are local, the electronic wave functions change slowly, and that the NACMEs are antisymmetric, $\vec{d}_{\beta\beta'}(\vec{R}_\mu) = -\vec{d}_{\beta'\beta}(\vec{R}_\mu)$. This is identical to Eq. (15.13) derived in the (simple) Ehrenfest method.

15.3.3 Computational Aspects

Standard practice in the simulations is to sample the phase space of the ground state using a Wigner distribution and allocate the initial coordinates to electronic states according to the transition dipole moments and excitation energy. This approximation, although common, does not explicitly account for the excitation process and there are now procedures to do so even in trajectory based simulations [48, 59, 82, 95]. In low-dimensionality systems where numerically exact quantum propagation methods are applicable, this is common procedure [4, 117] and constitutes a prerequisite for many types of coherent control schemes [3, 102]. For a calculation that starts in the ground state of the system, it is generally easiest to obtain the initial ground state by imaginary time propagation or by solving for the lowest eigenvalue of the Hamiltonian matrix in the basis of the ansatz wave functions [48]. For simulations of molecules, Cartesian coordinates are most commonly used, although other effective or internal coordinates are of course possible. It is important to evaluate the level of electronic structure theory used to calculate energies, gradients, and couplings to ensure that it is appropriate for the dynamics one wishes to simulate [8]. Examples of how simulations can be analysed are provided via the examples in the Application Section 15.4. Furthermore, one should examine how the simulations converge with respect to the number of trajectories, and also with respect to the sampling methods employed as discussed next.

The approximate form of propagation employed in MCE presents advantages in terms of numerical convenience, scalability, and stability compared to fully variational approaches such as vMCG [77, 119, 120] and G-MCTDH [13], but as a consequence we must pay extra attention to convergence. The approximate nature of the propagation of the individual trajectories can be compensated for via improved sampling. This is an extensive topic, discussed at length in for instance Refs. [60, 87, 97]. Here, we restrict ourselves to pointing out that the propagation of swarms of trajectories ensures that overlap, and thus coupling, of the propagated coherent states is maintained.

A swarm is generated from a tight distribution (e.g., a Husimi or Wigner distribution) with small shifts in initial positions and momenta. The utility of each trajectory can be further improved by propagating a train of coherent states along each trajectory, rather than just one coherent state, which does not add much to the computational overheads of the calculations.

Another issue to consider is a built-in deficiency in the Ehrenfest ansatz. The Ehrenfest basis is guided by a mean-field potential, which can be advantageous when quantum transitions are frequent and there are dense manifolds of electronic states with similar topology. However, if the electronic states are quite different, quantum mechanics allows for the nuclear wave packets on the different states to propagate away from each other. The enforced sharing of the coherent state across electronic states in the Ehrenfest wave packet prevents this, with an associated risk of over-coherence and misguiding of the Ehrenfest trajectories. There is a similar problem of over-coherence in surface-hopping because each surface-hopping trajectory also carries coefficients for all states (see e.g., [94]).

The situation can be remedied by a procedure included in *Ab-Initio* Multiple Cloning (AIMC-MCE) algorithm [57], inspired by *Ab-Initio* Multiple Spawning (AIMS) [62, 64]. The procedure is applied shortly after a trajectory passes near a conical intersection, when the non-adiabatic coupling is sufficiently low and if the "breaking force" that pulls the state β away from the other remaining states,

$$F_{\beta\mu}^{\text{break}} = |a_\mu^\beta|^2 \left(|\vec{\nabla}_I E_\beta(\vec{R}_\mu)| - \sum_{\beta'} |a_\mu^{\beta'}|^2 |\vec{\nabla}_I E_{\beta'}(\vec{R}_\mu)| \right), \tag{15.45}$$

is sufficiently strong. The procedure clones the Ehrenfest wave packet to render two clones, where one proceeds on the single potential energy surface β while the second is guided by an Ehrenfest force on the remaining electronic states. Although this procedure means that the number of trajectories increases during the simulation, it does act to remedy the single-set limitation of the MCE method. In a similar spirit, numerical procedures can be applied to improve the tunneling of the classically guided Ehrenfest wave packets [61].

15.4 Applications

In this section, we present illustrations of the theoretical dynamics methods just reviewed to simulate electron and nuclear dynamics of medium-sized organic molecules (Figure 15.1). In the first example, the (simple) Ehrenfest method is used to simulate coupled electron and nuclear dynamics upon sudden ionization, a topic of great interest with the recent development of attoscience. In particular, we study the effect of the nuclear motion on coherent electron dynamics and the nature of

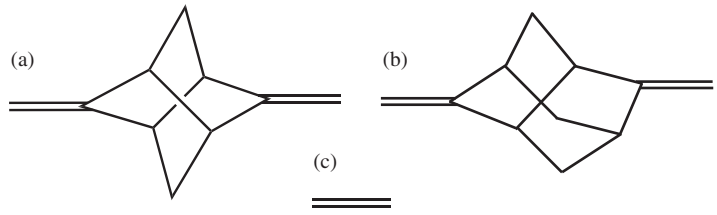


Figure 15.1 Structure of the molecules studied in the present applications. (a) BMA[5,5] and (b) BMA[6,5]: modified bismethylene-adamantane molecules where the cage consists of four connected cyclopentane rings, and two cyclohexane and two cyclopentane rings respectively. (c) Ethylene.

nuclear motion induced by an electronic wave packet. In the second example, inspired by the recent and important developments of new pulsed x-ray and electron sources capable of imaging molecular motion with femtosecond resolution, the MCE method is used to simulate non-adiabatic nuclear dynamics and to predict the x-ray and electron scattering signals that could be expected to be seen in an experiment. Importantly, the simulations allow us to investigate the effect of dispersion and non-locality in the wave packet on the experimental signal, and thus gain deeper understanding of what aspects of molecular motion can be observed in experiments.

15.4.1 Coupled Electron and Nuclear Dynamics Upon Sudden Ionization

Because of the time-energy uncertainty principle $\Delta E \Delta t \geq \hbar$, attosecond pulses have a large spectral bandwidth and ionization by them leads to the coherent population of several cationic electronic states, thus breaking the Born–Oppenheimer approximation. The system is no longer confined to being in a single stationary electronic state but is now a superposition of electronic states, called an electronic wave packet: such a superposition is non-stationary, i.e., its probability density is time-dependent. This type of pure electron dynamics is called "charge migration" in the literature [18] and can happen without nuclear motion.

Initial theoretical studies of charge migration considered only the electronic degrees of freedom [76]: typically, a single and fixed nuclear geometry was used. This means neglecting both the motion of the nuclei and the spreading of the nuclear wave packet. In such a framework, it was predicted that quantum interference between the populated electronic eigenstates would alternate between constructive and destructive, and thus leads to long-lived oscillating motion of the electronic density with a period inversely proportional to the electronic energy gap. A fundamental challenge has been to understand to what extent the electronic wave packet retains its coherence, i.e., how long the oscillations in the electronic density survive in the presence of interactions with the nuclear degrees of freedom. We argue that the mean-field "single-set" class of methods (Ehrenfest being the simplest) is particularly natural for the simulation of such coupled electron–nuclear dynamics since each nuclear trajectory "feels" the multiple coupled potential energy surfaces.

In this example, we present results of a theoretical study [112] of coupled electron and nuclear dynamics in modified bis-methylene adamantanes (BMA) (Figures 15.1a and b) using the Ehrenfest method. The two lowest-energy electronic states of BMA cation correspond to an electron being removed from the π bonds. The equilibrium geometry of the neutral BMA is a point of degeneracy between these two cationic states and thus, there is no pure electron dynamics upon valence ionization of the π system in this molecule (since the period of oscillation in the electronic density is inversely proportional to the energy gap). An appropriate chemical modification can be used to lift the degeneracy and "engineer" electron dynamics [110]. Here, we consider two modified molecular species where the number of carbon atoms in the cage rings are altered: BMA[6,5] where two of the cyclohexane rings of the cage have been replaced by cyclopentanes (Figure 15.1(b)) and BMA[5,5] where all four rings are cyclopentanes (Figure 15.1(a)).

The effect of the mean-field nuclear motion was investigated by comparing simulations of pure electron dynamics and simulations of electron dynamics coupled to Ehrenfest nuclear motion (considering a single nuclear geometry to start with). For the electronic structure, the complete active space self-consistent field (CASSCF) method is used, state-averaging over the two lowest-energy states, with the 4 π orbitals as active; the 6-31G* basis set is used [78]. The equations of motion are integrated with a fixed mass-weighted length step size of 0.01 $\text{amu}^{1/2}$bohr (corresponding to a varying time step of approximately 0.1 fs). To follow the evolution of the electronic wave packet, the electronic spin density – that locates the unpaired electron – is computed and partitioned on

the atoms [70] at each step of the simulation. The implementation of the above methods in a development version of Gaussian was used [27]. Figures 15.2a and b show the time evolution of the spin density on the left methylene group upon population of an equal in-phase superposition of the two lowest-energy states: $\Psi = \frac{1}{\sqrt{2}}(\Psi_0 + \Psi_1)$. With fixed nuclei (solid lines), the unpaired electron initially localized on the left methylene group migrates to the other methylene group, and then returns to its initial position. The whole process repeats itself and the hole thus oscillates back and forth between the two terminal bonds. The nature and amplitude of the electron dynamics are the same for the two molecules. However, the time scale is different, as expected, because of the different energy gap between the two electronic states. The effect of the nuclear motion on electron dynamics (dashed lines) is different in the two molecules. In BMA[5,5], the electron dynamics is barely affected by the nuclear motion. In BMA[6,5], the electron dynamics with nuclei fixed and moving nuclei are identical up to approximately 5 fs; then, the oscillations get faster when the nuclei are allowed to move. Also, the unpaired electron does not come back completely to its initial position since the amplitude of the oscillations decreases. In brief, the nuclear motion can alter both the period and amplitude of the oscillations in the electronic density from only a few femtoseconds. The magnitude of the effect of the nuclear motion is however very system-dependent [107, 110]. The smaller the energy gap, the slower the electron dynamics, the stronger the coupling with the nuclear coordinates and the more significant the effect of the nuclear motion on the electron dynamics.

The effect of the width of the nuclear wave packet on the electron dynamics was studied in BMA[5,5] with simulations started at 500 nuclear geometries and velocities sampled from a Wigner distribution to mimic the vibrational ground state nuclear wave packet of the neutral species before ionization. Figure 15.2(c) shows again the time evolution of spin density on the left methylene group, but for the ensemble of trajectories simulated independently: the individual oscillations dephase with time because of the different energy gaps at the different sampled geometries. The solid lines show the average spin density oscillation amplitude for different sample sizes. It is important to note that the result depends on the number of samples taken into account. Care must be taken so that convergence is reached since a too small ensemble may lead to the misleading result that electron dynamics would survive longer. With an ensemble of 500 geometries, the mean oscillation is quickly damped with time, resulting in the spin density being, on average, equally delocalized over the two methylene groups. A coherence half-life of $\approx$8 fs is obtained. In short, the effect of the nuclear wave packet width is substantial: it leads to a fast dephasing of the electron dynamics [42, 43, 111]. In BMA[5,5], it is larger than the effect of the nuclear motion per se. Because of the relatively "long" coherence half-life (8 fs) and because of the short period of oscillations (4.9 fs), several oscillations in the electronic density can be observed before dephasing occurs.

How are the nuclei actually moving in the meantime? Nuclear motion induced by a coherent superposition of electronic states is called "charge-directed reactivity" in the literature, suggesting that non-stationary electronic wave packets could play a chemical role. We present here an analysis of the nuclear motion in BMA[5,5] upon population of the equal in-phase superposition of the two lowest-energy states of the cation. As a reference, population of the ground state of the cation solely leads to symmetric stretching of the two methylene bonds with a period of about 20 fs, as expected since an electron has been removed (symmetrically) from the two bonding π bonds. Figure 15.3(a) shows the time evolution of the average (blue) and difference (cyan) in bond lengths of the two methylene groups in the case of the non-stationary electronic wave packet, for a single nuclear trajectory. The superposition of the two lowest-energy electronic states induces an asymmetric stretching of the two terminal methylene bonds because of the unpaired electron not being equally delocalized over the two terminal groups. The oscillations with time in the electronic

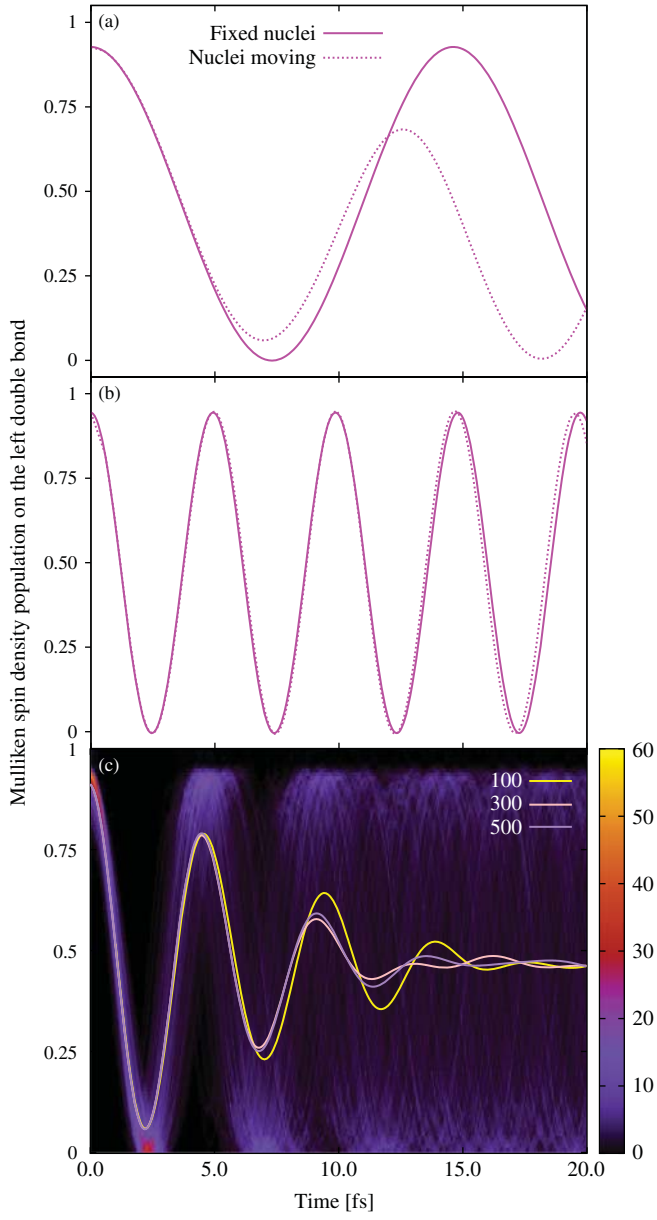


Figure 15.2 Time evolution of the spin density on the left methylene group in (a) BMA[6,5], (b,c) BMA[5,5]. (a,b) Simulations with a single nuclear geometry with fixed nuclei (solid) and nuclei moving (dashed). (c) Simulation for an ensemble of 500 geometries. The colour axis indicates the number of sampled trajectories in each pixel of the heat map. The solid lines indicate the spin density averaged over ensembles of different sizes. Adapted from Ref. [112] with permission from The Royal Society of Chemistry.

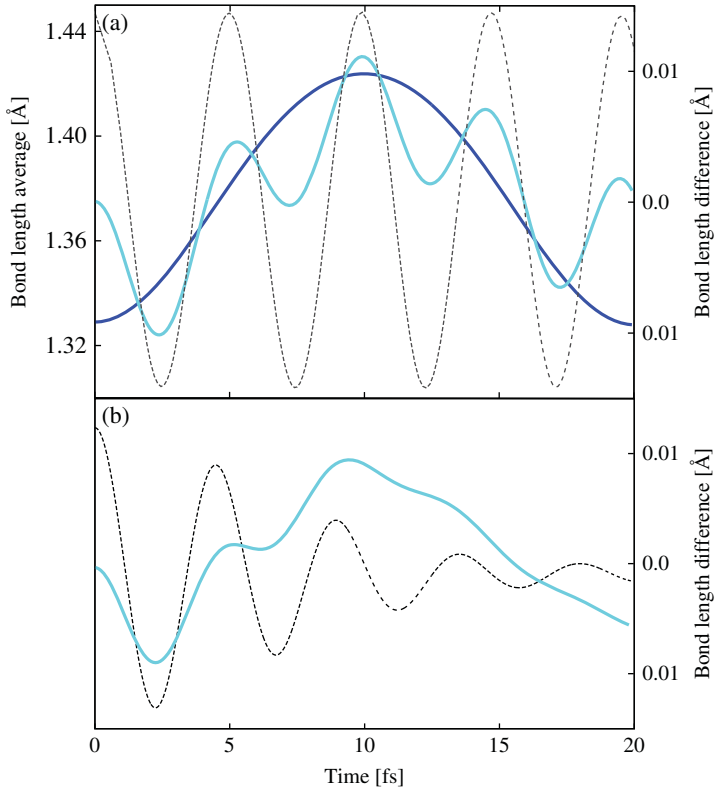


Figure 15.3 Time evolution of the bond lengths of the two methylene groups (average length in blue with the left *y*-axis and difference in cyan with the right *y*-axis) in BMA[5,5] for (a) a single nuclear trajectory and (b) an ensemble of 500 trajectories (with sampled geometries and velocities). The black dashed lines remind the reader of the oscillations in the electronic density from Figures 15.2b and c. Adapted from Ref. [112] with permission from The Royal Society of Chemistry.

density result in oscillations in the nuclear motion as well, the nuclei continuously trying to adapt to the time-dependent electronic distribution. Figure 15.3(b) shows the time evolution of the difference in bond lengths of the two methylene groups, averaged over the ensemble of 500 trajectories to take into account the width of the nuclear wave packet. The oscillations in the bond length difference are damped just as the oscillations in the electronic density dephase (recalled by the black dashed curve). In short, this example and others [66, 109] illustrate the "charge-directed reactivity" concept but also its limitation. The effect of the nuclear wave packet width is again crucial. As the electron dynamics dephase and the unpaired electron becomes equally delocalized over the two methylene groups, the asymmetry in the stretching of the two π bonds disappears.

To conclude: we have presented theoretical simulations of electron dynamics coupled to nuclear motion treated classically with the Ehrenfest method and where the intrinsic distribution of geometries in a nuclear wave packet is taken into account using an ensemble of independent trajectories. Although these are obviously a more realistic description than simulations considering only the electronic degrees of freedom, one can question the approximations used. Indeed, the effect of the nuclear motion may be underestimated with a mean-field approach and the quantum behavior of the nuclei will not be described by independent trajectories. One of the main drawbacks of the

(simple) Ehrenfest method is that the nuclear wave packets on different electronic states cannot move in different ways, which could affect the electronic coherence and therefore the predicted electron dynamics. Recent theoretical studies of "charge migration" coupled to quantum nuclear motion using the DD-vMCG method confirm the ultrafast electronic decoherence due to a dephasing mechanism [114].

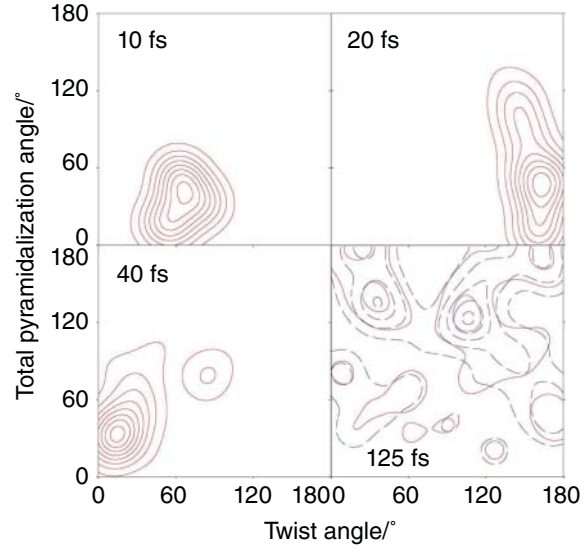
15.4.2 Ultrafast Scattering as a Probe of Nuclear Dynamics

In recent years, rapid development of pulsed x-ray and electron sources have made ultrafast scattering experiments possible [12, 29, 45, 53, 67–69, 79, 81, 88, 91, 116, 121, 122, 124]. Scattering methods probe the structure of molecules and sequences of time-resolved structures can then be assembled into so-called *molecular movies* that show the structural dynamics of the molecule. The complementarity between spectroscopy, which predominantly probe molecules in the energy domain, and scattering means that powerful insight into complex photo-chemical processes can be gained by combining observations from the two types of techniques [46, 75]. In this section, we will examine ultrafast x-ray and electron scattering using simulations of the molecule ethylene shown in Figure 15.1(c), and consider the effect of the nuclear wave packet, its delocalization and dispersion, on the experimental signal, as well as the importance of simulation parameters such as the wave packet widths [49, 50, 92].

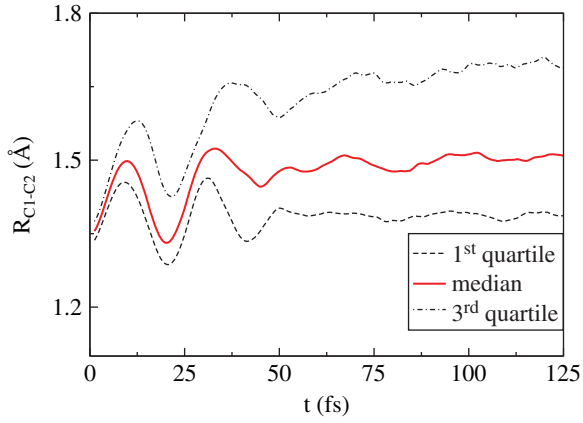
The non-adiabatic dynamics of ethylene following excitation by an optical pump laser is simulated using the AI-MCE approach [50, 80, 92]. The electronic potential energies, their gradients, and the non-adiabatic couplings are calculated *on-the-fly* at the three-state-averaged CASSCF level using the MOLPRO quantum chemistry package [115], with a small CAS(2,2) active space, known to describe the lowest two electronic excited states qualitatively [9, 54, 80], and the Dunning cc-ppVDZ basis [101]. The initial conditions are obtained using Monte Carlo sampling of the Wigner distribution [11] for the $v = 0$ vibrational ground state. The initial electronic population is assumed to be completely localized on the first excited S_1 $\pi\pi^*$ state. A total of 1000 Ehrenfest trajectories are propagated for 150 fs with a 0.1 fs time step.

The simulations show that upon photo-excitation to the Franck–Condon region of the $\pi\pi^*$ electronic S_1 state, ethylene undergoes cis-trans isomerization around the C=C double bond followed by decay via two competing processes. The first is non-radiative decay through a twisted or pyramidalized conical intersection, and the second is H-atom migration to form ethylidene (CH_3CH), which then decays through a different conical intersection (although CAS(2,2) does not account for H-migration very accurately [80]). The population of S_1 changes slowly during the first 30 fs, at which point the excited molecular wave function reaches a region where the gap between S_1 and S_0 is sufficiently small to allow for efficient population transfer and the population then decays exponentially with an approximate lifetime of $\tau \approx 112$ fs. The overall dynamics is consistent with previous calculations using AIMS [54].

The time evolution of the nuclear wave function is depicted in Figure 15.4(a) in terms of a probability density contour plot in the twist and pyramidalization angles. The twist angle corresponds to a rotation around the carbon–carbon double bond, and the pyramidalization angle reflects the degree of deviation of the two carbons from sp^2 hybridization. At first, the nuclear wave function moves ballistically (i.e., without spreading significantly) along the twisting coordinate with a period of ∼40 fs (the twisting of ethylene on the electronic ground state S_0 has a period of ∼33 fs). During the first twisting cycle, the wave function is almost totally located on the S_1 state but at later times the population starts transferring to the S_0 state. From 50 fs and onwards, the wave function



(a) Probability density



(b) C–C bond distance

Figure 15.4 (a) Contour plot showing the probability distribution of the C=C twist angle and the degree of pyramidalization of the carbon atoms at times $t = 10$, 20, 40, and 125 fs (with $t = 0$ set by the pump pulse). The solid red lines show the probability contour for the wave packet on the S_1 state, while the dashed blue line (which appears in the 125 fs frame) corresponds to the probability contour on the S_0 state. (b) Median distance between the carbon atoms in Å as a function of time in fs (solid red line), calculated for 1000 trajectories. The first (black dashed) and third quartile (black dot-dashed) are also shown. Adapted from Ref. [50] under CC-BY license from the American Chemical Society.

becomes increasingly dispersed. The dynamics can also be viewed in terms of the C-C bond distance as shown in Figure 15.4(b), which shows the median C-C bond distance, as well as the first and third quartiles which together enclose 50% of the bond distances. At short times, the C-C bond oscillates coherently, with the dispersion becoming strongly apparent after approximately 50 fs as previously seen in the density plot.

We proceed to calculate the scattering signal from the ethylene simulation, using the well-established independent atom model (IAM) [1, 20]. The IAM has shortcomings [15, 72, 73], but provides more than sufficient accuracy in the present context. According to IAM the

rotationally averaged elastic scattering is proportional to the squared molecular form factor $|f_{\mathrm{IAM}}(q,\tau)|^2$ given by,

$$|f_{\mathrm{IAM}}(q,\tau)|^2 = \sum_{I=1}^{N_{\mathrm{at}}} |f_I^0(q)|^2 + \sum_{J\neq I}^{N_{\mathrm{at}}} f_I^0(q) f_J^0(q) \frac{\sin(qR_{IJ}(\tau))}{qR_{IJ}(\tau)}, \tag{15.46}$$

where $R_{IJ}(\tau)$ is the distance between atoms I and J in the molecule, τ is the time, and $q = (4\pi/\lambda)\sin(\theta/2)$ is the amplitude of the momentum transfer vector for the scattering, with θ the deflection angle of the scattered particle relative to the direction of the incoming probe beam and λ its de Broglie wavelength. The atomic form factors ${f_I}^0(q)$ are tabulated for both x-ray and electron scattering [1]. In the case of electron scattering, they contain additional contributions to account for the scattering from nuclei as well as corrections to compensate for relativistic effects and deviations from the Born approximation [1]. The first sum on the right-hand side of Eq. (15.46) is the atomic term, which is independent of molecular structure, while the second sum is the molecular term, which contains the scattering interferences that yield structural information.

The differential scattering cross-section is proportional to the molecular scattering form factor convoluted by the molecular geometry, as given by the nuclear wave function $\Xi(\vec{R},\tau)$ and the intensity profile of the probe pulse $I(\tau)$ [49, 50],

$$\mathrm{d}S(q,t) = \int I(\tau)\, \langle \Xi(\vec{R},\tau)|\ |f_{\mathrm{IAM}}(q,\tau)|^2\ |\Xi(\vec{R},\tau)\rangle_{\vec{R}}\ \mathrm{d}\tau, \tag{15.47}$$

where t is the time-delay between the pump and probe pulses and the distances between atoms in the molecular form factor are seen to act as operators.

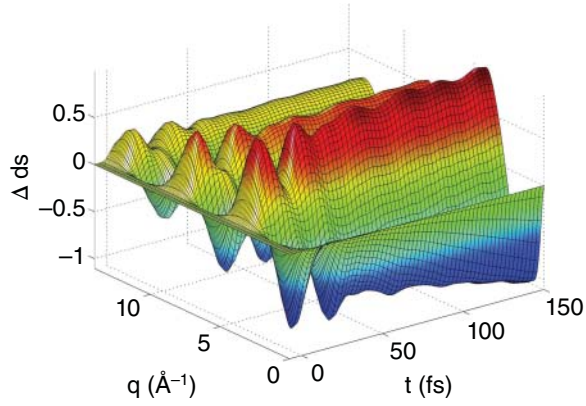
Finally, the diffraction signal $\Delta\mathrm{d}S(q,t)$ is given in terms of a *laser on–laser off* percentage difference signal, commensurate with how experimental data is presented [69],

$$\Delta\mathrm{d}S(q,t) = \gamma_{\mathrm{exc}} \frac{\mathrm{d}S(q,t) - \mathrm{d}S_{\mathrm{off}}(q)}{\mathrm{d}S_{\mathrm{off}}(q)}, \tag{15.48}$$

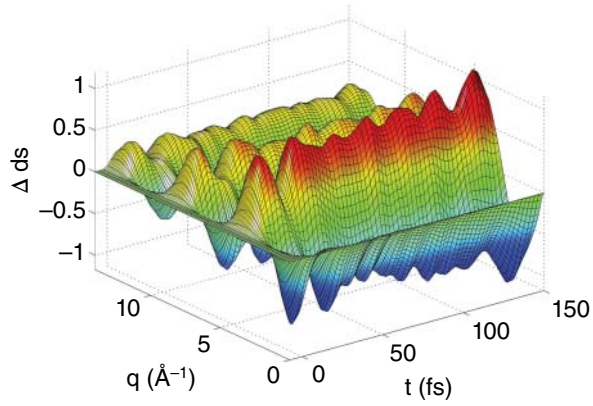
where γ_{exc} is the fraction of excited molecules, $\mathrm{d}S(q,t)$ is the *laser on* signal corresponding to the excited molecular wave function, and $\mathrm{d}S_{\mathrm{off}}(q)$ is the *laser off* background signal from unpumped molecules. A Gaussian-profile probe pulse with duration 25 fs (FWHM) is used, with the range of q running from 0 to 14 Å^{-1} (corresponding to 13.8 keV x-ray photons) and an excitation fraction for the molecules of $\gamma_{\mathrm{exc}} = 9\%$.

The rotationally averaged elastic x-ray scattering difference signal calculated according to Eq. (15.48) from the full simulation is shown in Figure 15.5 as a function of time. In IAM, the contribution to the signal from each atom is proportional to the number of electrons in the atom squared. The signal is therefore dominated by the two carbon atoms, which account for 12 out of the 16 electrons in the molecule, making the contribution from the four hydrogen atoms comparatively minor. The variation in the signal in Figure 15.5(a) becomes progressively smaller over time. This is to be anticipated given the increase in dispersion of the nuclear wave function observed in Figure 15.4, with the associated delocalization averaging out specific molecular geometries and motions. At times $t < 50$ fs, however, the coherent stretch of the C–C bond, clearly visible in Figure 15.4(b), results in a strong signature in the scattering signal across the whole range of the momentum transfer q.

The sensitivity of the calculated scattering signal $\Delta\mathrm{d}S(q,t)$ to the number of Ehrenfest trajectories included can be seen when comparing Figure 15.5(a) with Figure 15.5(b), which has been calculated using a small randomly selected subset of 20 trajectories rather than 1000. Although the exact appearance of the signal from the smaller set will depend on which 20 trajectories are included (meaning that the signal is prone to noise), a subset of 20 is sufficiently large to make these differences rather small. Interestingly, even the small subset correctly depicts the main features of the



(a) $N = 1000$ trajectories



(b) $N = 20$ trajectories

Figure 15.5 Elastic x-ray scattering difference signal, $\Delta \mathrm{d}S(q,t)$, in percent, shown as a function of momentum transfer, q in Å^{-1}, and pump-probe delay time, t in fs, calculated using Eq. (15.48) and rotational averaging. (a) Signal calculated for 1000 trajectories and (b) a representative signal for a subset of 20 randomly selected trajectories. Adapted from Ref. [50] under CC-BY license from the American Chemical Society.

scattering signal. This is consistent with previous findings [68, 104], which demonstrated that if the trajectories can be clustered into self-similar sets, then the experimental signal from an ultrafast scattering experiment, which is sensitive to molecular geometry above all, can be reproduced using only a comparatively small number of trajectories with the number corresponding to the number of self-similar sets. However, the smaller subset will generally (and does so here) underestimate the dispersion of the nuclear wave packet at longer times, as can be seen from the features present at long times in the small set in Figure 15.5(b), but absent from the full set in Figure 15.5(a). Thus using a single trajectory to represent the dynamics would strongly overestimate the amount of detail present in the signal.

We continue by examining the effect of the width of the individual wave packets. The widths are determined by the factors α of each coherent state as shown in Eq. (15.22). In Figure 15.6 we compare the default scattering difference signal, $\Delta \mathrm{d}S(q,t)$, with the signal obtained using modified values of α. Specifically, the comparison is to $\alpha/3$, corresponding to a more *delocalized*, and to 3α, corresponding to a more *localized* wave packet than default. The results show that the more

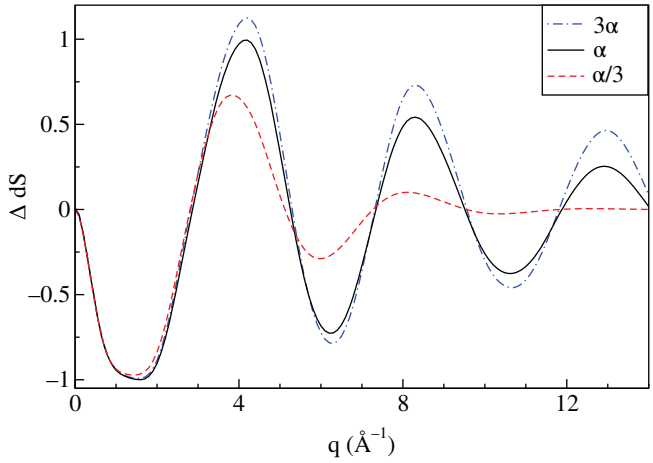


Figure 15.6 The percentage difference signal, $\Delta \mathrm{d}S(q,t)$, calculated according to Eq. (15.48) as a function of q in Å^{-1} at time $t = 100$ fs, using rotational averaging and 100 trajectories. The three curves are calculated identically except that the width parameters $\alpha = \{\alpha_C, \alpha_H\}$ have been manipulated a posteriori to examine the effect of α on the signal. Results are shown for the original values of α (black solid line), for less localized wave packets with $\frac{1}{3}\alpha$ (red dashed line), and for more localized wave packets with 3α (blue dot-dashed line). Adapted from Ref. [50] under CC-BY license from the American Chemical Society.

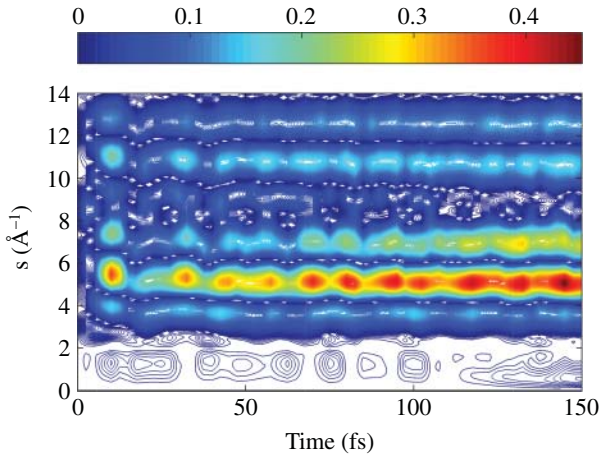


Figure 15.7 Comparison of the elastic scattering signal for photo-excited ethylene obtained with electron scattering versus x-ray scattering, plotted in a contour plot as a function of the momentum transfer q (s) and delay-time t. The absolute difference between the two, $|\Delta \mathrm{d}S^{\mathrm{elec}}(q,t) - \Delta \mathrm{d}S^{\mathrm{xray}}(q,t)|$, is shown. Adapted from Ref. [92] under CC-BY license from Elsevier.

delocalized the wave packet is, the weaker the signal becomes at large q. Even the comparatively modest reduction to $\alpha/3$ is sufficient to almost completely eliminate the signal for $q > 8$ Å^{-1}. It must be noted that due to the local phase-space coordinate propagation of the Ehrenfest wave packets, the AI-MCE quantum dynamics approach is robust only if sufficiently narrow coherent states are used as a basis for the nuclear dynamics [50]. Another important point is that if the quantum molecular dynamics simulations are fully converged, an admittedly rare situation, the total molecular wave function will have a "shape" which is independent of the α width parameters used.

The capacity of x-ray scattering to detect small features in the nuclear probability density is ultimately limited by the fact that it uses approximately atom-sized probes to sketch the nuclear probability density. What if we had a sharper, more point-like probe? This is exactly what happens in ultrafast electron diffraction. Electrons scatter not just from the electron density, but also from the nuclei, which allows electron diffraction to trace the nuclear probability distribution more sharply. In Figure 15.7, we compare the signal resulting from x-ray scattering with the signal from electron diffraction by showing $|\Delta \mathrm{d}S^{\mathrm{elec}}(q,t) - \Delta \mathrm{d}S^{\mathrm{xray}}(q,t)|$. The difference between the two signals is small at small values of q (corresponding to large distances in the molecule where the size of the probe is unimportant), but increases at intermediate and large values of q. The large values of q map onto small distances in the molecule, and the result thus confirms that the electron diffraction is able to probe smaller, more local features in the nuclear probability density. The electron scattering can also be shown to be more sensitive to the hydrogens in ethylene, and especially at longer times the most distinct nuclear motion is associated with the hydrogen atoms [92].

In conclusion, a detailed MCE simulation of a pump-probe experiment in ethylene has been used to examine how experimental observables relate to the simulation. We find that a description of the nuclear dynamics beyond a single trajectory is important in order to predict realistic signals, and that simulation parameters such as the width of the coherent states can have a strong influence on the results. As we approach a point where experiments probe the nuclear wave function in greater detail, the comparisons between experiment and theory will become more demanding. Methods such as MCE, which carry a reasonable (or at least feasible) computational cost and are, at least in principle, capable of representing the evolution of the nuclear wave function accurately, will become increasingly important. At the same time, this new generation of ultrafast experiments will provide opportunities to examine the veracity of simulations. In the context of scattering and diffraction, the physical model used to calculate the signals can be markedly improved by considering corrections to the independent atom model [15, 71–73], accounting for inelastic scattering [14, 16], and coherence effects [25, 35, 50, 90, 99]. The latter effects might even prove useful to observe electron dynamics using x-ray scattering [25, 48, 51, 82, 89, 90, 96]. With the rapid and ongoing development of simulations techniques and ultrafast experiments we can expect the intense interaction between experiments, simulations, and theory to continue.

15.5 Conclusion

In this chapter, we have reviewed possible theoretical methods for the description of coupled electron and nuclear dynamics. Both the Ehrenfest and MCE treat the electronic degrees of freedom quantum mechanically. To describe the nuclear degrees of freedom, the Ehrenfest method uses an ensemble of independent trajectories propagated classically on a mean-field potential, following the gradient of a superposition of electronic states. The main approximation of the Ehrenfest method is thus the mean-field treatment of the coupling between electronic and nuclear degrees of freedom. The MCE method extends this approach by using the ensemble of Ehrenfest trajectories as a basis to solve the time-dependent Schrödinger equation. The time evolution of the coupling coefficients of the Ehrenfest wave packets, and thus the transfer of population between them, is thus determined quantum mechanically, while individual Ehrenfest wave packet propagates classically, simplifying the computational implementation for direct dynamics. A practical approach for describing the nuclear wave packet branching properly is proposed. After deriving the equations of motion of these two methods, we have presented applications for each method illustrating their use

in simulating non-adiabatic dynamics of medium-sized organic molecules relevant to attosecond and femtosecond time-resolved experiments.

The rapid advances in ultrafast imaging techniques will undoubtedly continue, enabling ever more detailed comparisons between experiments and theory. In response to this challenge, it is important that existing simulation methods for non-adiabatic molecular quantum dynamics continue to be developed. This must include improvements in the ab-initio electronic structure calculations and the propagation of the nuclear wave packet, but also in methods to calculate experimental observables such as photo-ionization or scattering directly from the simulations. Eventually, the combination of new experiments and state-of-the-art theory should lead to a more detailed and complete understanding of both photo-chemistry and photo-physics.

References

1 *International Tables for Crystallography Volume C: Mathematical, Physical and Chemical Tables* (ed. Prince, E.), online edition (2006). International Union of Crystallography. ISBN: 978-1-4020-1900-5.

2 Kirrander, A. and Jungen, Ch. (2011). Molecular ion-pair states in ungerade H_2. *Phys. Rev. A* 84: 052512.

3 Kirrander, A., Jungen, Ch., and Fielding, H.H. (2008). Localization of electronic wave packets in H_2. *J. Phys. B* 41: 074022.

4 Kirrander, A., Jungen, Ch., and Fielding, H.H. (2010). Control of ionization and dissociation with optical pulse trains. *Phys. Chem. Chem. Phys.* 12 (31): 8948.

5 Abedi, A., Maitra, N.T., and Gross, E.K.U. (2010). Exact factorization of the time-dependent electron-nuclear wave function. *Phys. Rev. Lett.* 105: 123002.

6 Almlöf, J. and Taylor, P.R. (1985). Molecular properties from perturbation theory: a unified treatment of energy derivatives. *Int. J. Quantum Chem.* 27 (6): 743–768.

7 Amarouche, M., Gadea, F.X., and Durup, J. (1989). A proposal for the theoretical treatment of multi-electronic-state molecular dynamics: hemiquantal dynamics with the whole dim basis (HWD). A test on the evolution of excited Ar_3^+ cluster ions. *Chem. Phys.* 130 (1–3): 145–157.

8 Bellshaw, D., Minns, R.S., and Kirrander, A. (2019). Correspondence between electronic structure calculations and simulations: nonadiabatic dynamics in CS_2. *Phys. Chem. Chem. Phys.* 21 (26): 14226–14237.

9 Ben-Nun, M. and Martinez, T.J. (2000). Photodynamics of ethylene: *ab initio* studies of conical intersections. *Chem. Phys.* 259 (1–3): 237–248.

10 Born, M. and Huang, K. (1968). *Dynamical Properties of Crystal Lattices*. Oxford: Clarendon. ISBN: 9780198503699.

11 Brown, R.C. and Heller, E.J. (1981). Classical trajectory approach to photodissociation: the Wigner method. *J. Chem. Phys.* 75: 186.

12 Budarz, J.M., Minitti, M.P., Cofer-Shabica, D.V. et al. (2016). Observation of femtosecond molecular dynamics via pump-probe gas phase Xray scattering. *J. Phys. B* 49 (3): 034001.

13 Burghardt, I., Meyer, H.-D., and Cederbaum, L.S. (1999). Approaches to the approximate treatment of complex molecular systems by the multiconfiguration time-dependent Hartree method. *J. Chem. Phys.* 111 (7): 2927–2939.

14 Moreno Carrascosa, A. and Kirrander, A. (2017). *Ab initio* calculation of inelastic scattering. *Phys. Chem. Chem. Phys.* 19 (30): 19545–19553.

15 Moreno Carrascosa, A., Northey, T., and Kirrander, A. (2017). Imaging rotations and vibrations in polyatomic molecules with X-ray scattering. *Phys. Chem. Chem. Phys.* 19 (11): 7853–7863.

16 Moreno Carrascosa, A., Yong, H., Crittenden, D.L. et al. (2019). Ab-initio calculation of total X-ray scattering from molecules. *J. Chem. Theory Comp.* 15: 2836–2846.

17 Cederbaum, L.S. (2013). The exact molecular wave function as a product of an electronic and a nuclear wave function. *J. Chem. Phys.* 138 (22): 224110.

18 Cederbaum, L.S. and Zobeley, J. (1999). Ultrafast charge migration by electron correlation. *Chem. Phys. Lett.* 307 (3–4): 205–210.

19 Christensen, L., Nielsen, J.H., Brandt, C.B. et al. (2014). Dynamic stark control of torsional motion by a pair of laser pulses. *Phys. Rev. Lett.* 113: 073005.

20 Debye, P. (1915). Zerstreuung von Röntgenstrahlen. *Ann. Phys.* 46: 809.

21 Delos, J.B. and Thorson, W.R. (1979). Diabatic and adiabatic representations for atomic collision processes. *J. Chem. Phys.* 70 (4): 1774–1790.

22 Delos, J.B. and Thorson, W.R. (1972). Semiclassical theory of inelastic collisions. II. Momentum-space formulation. *Phys. Rev. A* 6: 720–727.

23 Delos, J.B., Thorson, W.R., and Knudson, S.K. (1972). Semiclassical theory of inelastic collisions. I. Classical picture and semiclassical formulation. *Phys. Rev. A* 6: 709–720.

24 Dirac, P.A.M. (1930). Note on exchange phenomena in the thomas atom. *Math. Proc. Camb. Philos. Soc.* 26, 7: 376–385.

25 Dixit, G., Vendrell, O., and Santra, R. (2012). Imaging electronic quantum motion with light. *Proc. Natl. Acad. Sci.* 109 (29): 11636.

26 Domcke, W. and Yarkony, D.R. (2012). Role of conical intersections in molecular spectroscopy and photoinduced chemical dynamics. *Annu. Rev. Phys. Chem.* 63: 325.

27 Frisch, M.J., Trucks, G.W., Schlegel, H.B. et al. (2010). Gaussian development version, revision h.32. Wallingford CT: Gaussian, Inc.

28 Gerber, R.B., Buch, V., and Ratner, M.A. (1982). Time-dependent self-consistent field approximation for intramolecular energy transfer. I. Formulation and application to dissociation of van der Waals molecules. *J. Chem. Phys.* 77 (6): 3022–3030.

29 Glownia, J.M., Natan, A., Cryan, J.P. et al. (2016). Self-referenced coherent diffraction X-ray movie of Ångstrom-and femtosecond-scale atomic motion. *Phys. Rev. Lett.* 117: 153003.

30 Goldstein, H. (1980). *Classical Mechanics*, 2e. Addison-Wesley.

31 Hack, M.D., Wensmann, A.M., Truhlar, D.G. et al. (2001). Comparison of full multiple spawning, trajectory surface hopping, and converged quantum mechanics for electronically nonadiabatic dynamics. *J. Chem. Phys.* 115 (3): 1172–1186.

32 Heller, E.J. (1976). Time dependent variational approach to semiclassical dynamics. *J. Chem. Phys.* 64 (1): 63–73.

33 Heller, E.J. (1981). Frozen Gaussians: a very simple semiclassical approximation. *J. Chem. Phys.* 75 (6): 2923.

34 Heller, E.J. (1981). The semiclassical way to molecular spectroscopy. *Acc. Chem. Res.* 14 (12): 368.

35 Henriksen, N.E. and Møller, K.B. (2008). On the theory of time-resolved X-ray diffraction. *J. Phys. Chem. B* 112: 558.

36 Hentschel, M., Kienberger, R., Spielmann, C. et al. (2001). Attosecond metrology. *Nature* 414 (6863): 509–513.

37 Herman, M.F. (1994). Dynamics by semiclassical methods. *Annu. Rev. Phys. Chem.* 45 (1): 83–111.

38 Hockett, P., Bisgaard, C.Z., Clarkin, O.J., and Stolow, A. (2011). Time-resolved imaging of purely valence-electron dynamics during a chemical reaction. *Nat. Phys.* 7 (8): 612–615.

39 Hunter, G. (1975). Conditional probability amplitudes in wave mechanics. *Int. J. Quantum Chem.* 9 (2): 237–242.

40 Ibrahim, H., Wales, B., Beaulieu, S. et al. (2014). Tabletop imaging of structural evolutions in chemical reactions demonstrated for the acetylene cation. *Nat. Commun.* 5: 4422.

41 Jasper, A.W., Nangia, S., Zhu, C., and Truhlar, D.G. (2006). Non-born-oppenheimer molecular dynamics. *Acc. Chem. Res.* 39 (2): 101–108.

42 Jenkins, A.J., Vacher, M., Bearpark, M.J., and Robb, M.A. (2016). Nuclear spatial delocalization silences electron density oscillations in 2-phenyl-ethyl-amine (PEA) and 2-phenylethyl-n,n-dimethylamine (PENNA) cations. *J. Chem. Phys.* 144 (10): 104110.

43 Jenkins, A.J., Vacher, M., Twidale, R.M. et al. (2016). Charge migration in polycyclic norbornadiene cations: winning the race against decoherence. *J. Chem. Phys.* 145 (16): 164103.

44 Jenkins, A.J., Spinlove, K.E., Vacher, M., Worth, G.A., and Robb, M.A. (2018). The Ehrenfest method with fully quantum nuclear motion (Qu-Eh): application to charge migration in radical cations. *J. Chem. Phys.*

45 Kim, K.H., Kim, J.G., Nozawa, S. et al. (2015). Direct observation of bond formation in solution with femtosecond X-ray scattering. *Nature* 518: 385.

46 Kim, K.H., Kim, J., Oang, K.Y. et al. (2015). Identifying the major intermediate species by combining time-resolved X-ray solution scattering and X-ray absorption spectroscopy. *Phys. Chem. Chem. Phys.* 17 (36): 23298–23302.

47 Kirrander, A. (2010). Heavy Rydberg states: the H^+H^-system. *J. Chem. Phys.* 133: 121103.

48 Kirrander, A. and Shalashilin, D.V. (2011). Quantum dynamics with fermion coupled coherent states. *Phys. Rev. A* 84: 033406.

49 Kirrander, A. and Weber, P.M. (2017). Fundamental limits on spatial resolution in ultrafast X-ray diffraction. *Appl. Sci.* 7 (6): 534.

50 Kirrander, A., Saita, K., and Shalashilin, D.V. (2016). Ultrafast X-ray scattering from molecules. *J. Chem. Theory Comput.* 12: 957–967.

51 Kowalewski, M., Bennett, K., and Mukamel, S. (2017). Monitoring nonadiabatic avoided crossing dynamics in molecules by ultrafast X-ray diffraction. *Struct. Dyn.* 4: 054101.

52 Kramer, P. and Saraceno, M. (1981). *Geometry of the Time-Dependent Variational Principle in Quantum Mechanics*, 1e. Springer Berlin Heidelberg. ISBN: 978-3-540-10579-4.

53 Levantino, M., Schiró, G., Lemke, H.T. et al. (2015). Ultrafast myoglobin structural dynamics observed with an X-ray free-electron laser. *Nature Comm.* 6: 6772.

54 Levine, B.G., Coe, J.D., Virshup, A.M., and Martinez, T.J. (2008). Implementation of *ab initio* multiple spawning in the MOLPRO quantum chemistry package. *Chem. Phys.* 347: 3.

55 Li, X., Tully, J.C., Schlegel, H.B., and Frisch, M.J. (2005). Ab initio Ehrenfest dynamics. *J. Chem. Phys.* 123 (8): 084106.

56 Lichten, W. (1963). Resonant charge exchange in atomic collisions. *Phys. Rev.* 131: 229–238.

57 Makhov, D.V., Glover, W.J., Martinez, T.J., and Shalashilin, D.V. (2014). *Ab initio* multiple cloning algorithm for quantum nonadiabatic molecular dynamics. *J. Chem. Phys.* 141: 054110.

58 Makhov, D.V., Saita, K., Martinez, T.J., and Shalashilin, D.V. (2015). *Ab initio* multiple cloning simulations of pyrrole photodissociation: TKER spectra and velocity map imaging. *Phys. Chem. Chem. Phys.* 17: 3316.

59 Makhov, D.V. and Shalashilin, D.V. (2018). Floque Hamiltonian for incorporating electronic excitation by a laser pulse into simulations of non-adiabatic dynamics. *Chem. Phys.*

60 Makhov, D.V., Martinez, T.J., and Shalashilin, D.V. (2016). Toward fully quantum modelling of ultrafast photodissociation imaging experiments. Treating tunnelling in the ab initio multiple cloning approach. *Faraday Discuss.*

61 Makhov, D.V., Symonds, C., Fernandez-Alberti, S., and Shalashilin, D.V. (2017). Ab initio quantum direct dynamics simulations of ultrafast photochemistry with multiconfigurational Ehrenfest approach. *Chem. Phys.* 493: 200–218.

62 Martinez, T., Ben-Nun, M., and Levine, R.D. (1996). Multi-electronic-state molecular dynamics: a wave function approach with applications. *J. Phys. Chem.* 100 (19): 7884–7895.

63 Martinez, T.J. (1997). Ab initio molecular dynamics around a conical intersection: $Li(2p)+H_2$. *Chem. Phys. Lett.* 272 (3–4): 139–147.

64 Martinez, T.J., Ben-Nun, M., and Ashkenazi, G. (1996). Classical/quantal method for multistate dynamics: a computational study. *J. Chem. Phys.* 104 (8): 2847–2856.

65 Mead, C.A. and Truhlar, D.G. (1982). Conditions for the definition of a strictly diabatic electronic basis for molecular systems. *J. Chem. Phys.* 77 (12): 6090–6098.

66 Meisner, J., Vacher, M., Bearpark, M.J., and Robb, M.A. (2015). Geometric rotation of the nuclear gradient at a conical intersection: extension to complex rotation of diabatic states. *J. Chem. Theory Comput.* 11 (7): 3115–3122.

67 Miller, R.J.D. (2016). Ultrafast imaging of photochemical dynamics: roadmap to a new conceptual basis for chemistry. *Faraday Discuss.* 194: 777–828.

68 Minitti, M.P., Budarz, J.M., Kirrander, A. et al. (2015). Imaging molecular motion: femtosecond X-ray scattering of an electrocyclic chemical reaction. *Phys. Rev. Lett.* 114 (25): 255501.

69 Minitti, M.P., Budarz, J.M., Kirrander, A. et al. (2014). Toward structural femtosecond chemical dynamics: imaging chemistry in space and time. *Faraday Discuss.* 171: 81.

70 Mulliken, R.S. (1955). Electronic population analysis on LACO-MO molecular wave functions. I. *J. Chem. Phys.* 23 (10): 1833–1840.

71 Northey, T. and Kirrander, A. (2019). Ab initio fragment method for calculating molecular X-ray diffraction. *J. Phys. Chem. A* 123 (15): 3395–3406.

72 Northey, T., Zotev, N., and Kirrander, A. (2014). *Ab Initio* calculation of molecular diffraction. *J. Chem. Theory Comput.* 10 (11): 4911.

73 Northey, T., Moreno Carrascosa, A., Schäfer, S., and Kirrander, A. (2016). Elastic X-ray scattering from state-selected molecules. *J. Chem. Phys.* 145: 154304.

74 Paul, P.M., Toma, E.S., Breger, P. et al. (2001). Observation of a train of attosecond pulses from high harmonic generation. *Science* 292 (5522): 1689–1692.

75 Pemberton, C.C., Zhang, Y., Saita, K., Kirrander, A., and Weber, P.M. (2015). From the (1B) spectroscopic state to the photochemical product of the ultrafast ring-opening of 1,3-Cyclohexadiene: a spectral observation of the complete reaction path. *J. Phys. Chem. A* 119 (33): 8832.

76 Remacle, F. and Levine, R.D. (2007). Probing ultrafast purely electronic charge migration in small peptides. *Z. Phys. Chem.* 221 (5): 647–661.

77 Richings, G.W., Polyak, I., Spinlove, K.E. et al. (2015). Quantum dynamics simulations using Gaussian wave packets: the vmcg method. *Int. Rev. Phys. Chem.* 34: 269–308.

78 Roos, B.O., Taylor, P.R., and Siegbahn, P.E.M. (1980). A complete active space SCF method (CASSF) using a density matrix formulated super-CI approach. *Chem. Phys.* 48 (2): 157–173.

79 Ruddock, J.M., Zotev, N., Stankus, B. et al. (2019). Simplicity beneath complexity: counting molecular electronsreveals transients and kinetics of photodissociation reactions. *Angew. Chem. Int. Ed.* 58 (19): 6371–6375.

80 Saita, K. and Shalashilin, D.V. (2012). On-the-fly *ab initio* molecular dynamics with multiconfigurational Ehrenfest method. *J. Chem. Phys.* 137: 22A506.

81 Sciaini, G. and Miller, R.J.D. (2011). Femtosecond electron diffraction: heralding the era of atomically resolved dynamics. *Rep. Prog. Phys.* 74: 096101.

82 Shalashilin, D.V., Child, M.S., and Kirrander, A. (2008). Mechanisms of double ionization in strong laser field from simulation with coupled coherent states. *Chem. Phys.* 347: 257.

83 Shalashilin, D.V. (2009). Quantum mechanics with the basis set guided by Ehrenfest trajectories: theory and application to spin-boson model. *J. Chem. Phys.* 130 (24): 244101.

84 Shalashilin, D.V. (2010). Nonadiabatic dynamics with the help of multiconfigurational Ehrenfest method: improved theory and fully quantum 24D simulation of pyrazine. *J. Chem. Phys.* 132 (24): 244111.

85 Shalashilin, D.V. (2011). Multiconfigurational Ehrenfest approach to quantum coherent dynamics in large molecular systems. *Faraday Discuss.* 153: 105.

86 Shalashilin, D.V. and Child, M.S. (2004). The phase space CCS approach to quantum and semiclassical molecular dynamics for high-dimensional systems. *Chem. Phys.* 304 (1–2): 103.

87 Shalashilin, D.V. and Child, M.S. (2008). Basis set sampling in the method of coupled coherent states: coherent state swarms, trains, and pancakes. *J. Chem. Phys.* 128 (5): 054102.

88 Shorokhov, D., Park, S.T., and Zewail, A.H. (2005). Ultrafast electron diffraction: dynamical structures on complex energy landscapes. *ChemPhysChem* 6: 2228.

89 Simmermacher, M., Henriksen, N.E., and Møller, K.B. (2017). Time-resolved X-ray scattering by electronic wave packets: analytic solutions to the hydrogen atom. *Phys. Chem. Chem. Phys.* 19: 19740–19749.

90 Simmermacher, M., Henriksen, N.E., Møller, K.B. et al. (2019). Electronic coherence in ultrafast X-ray scattering from molecular wave packets. *Phys. Rev. Lett.* 122: 073003.

91 Stankus, B., Budarz, J.M., Kirrander, A. et al. (2016). Femtosecond photodissociation dynamics of 1,4-diiodobenzene by gas-phase X-ray scattering and photoelectron spectroscopy. *Faraday Discuss.* 194: 525–536.

92 Stefanou, M., Saita, K., Shalashilin, D.V., and Kirrander, A. (2017). Comparison of ultrafast electron and X-ray diffraction – a computational study. *Chem. Phys. Lett.* 683: 300–305.

93 Stolow, A. (2013). The three pillars of photo-initiated quantum molecular dynamics. *Faraday Discuss.* 163: 9.

94 Subotnik, J.E. and Shenvi, N. (2011). A new approach to decoherence and momentum rescaling in the surface hopping algorithm. *J. Chem. Phys.* 134 (2): 024105.

95 Suchan, J., Hollas, D., Curchod, B.F.E., and Petr, S. (2018). On the importance of initial conditions for excited-state dynamics. *Faraday Discuss.* 212: 307–330.

96 Suominen, H.J. and Kirrander, A. (2014). How to observe coherent electron dynamics directly. *Phys. Rev. Lett.* 112: 043002.

97 Symonds, C., Kattirtzi, J.A., and Shalashilin, D.V. (2018). The effect of sampling techniques used in the multiconfigurational Ehrenfest method. *J. Chem. Phys.* 148 (18): 184113.

98 Takatsuka, K. and Yonehara, T. (2011). Exploring dynamical electron theory beyond the born–oppenheimer framework: from chemical reactivity to non-adiabatically coupled electronic and nuclear wave packets on-the-fly under laser field. *Phys. Chem. Chem. Phys.* 13 (11): 4987–5016.

99 Tanaka, S., Chernyak, V., and Mukamel, S. (2001). Time-resolved X-ray spectroscopies: Nonlinear response functions and liouville-space pathways. *Phys. Rev. A* 63: 063405.

100 Tannor, D.J. (2007). *Introduction to Quantum Mechanics: A Time-dependent Perspective*. University Science Books. ISBN: 978-1-891389-23-8.

101 Dunning, T.H. Jr., (1989). Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* 90: 1007.

102 Thomas, E.F. and Henriksen, N.E. (2017). Phase-modulated nonresonant laser pulses can selectively convert enantiomers in a racemic mixture. *J. Phys. Chem. Lett.* 8 (10): 2212–2219.

103 Thompson, A.L., Punwong, C., and Martinez, T.J. (2010). Optimization of width parameters for quantum dynamics with frozen Gaussian basis sets. *Chem. Phys.* 370: 70.

104 Tudorovskaya, M., Minns, R.S., and Kirrander, A. (2018). Effects of probe energy and competing pathways on time-resolved photoelectron spectroscopy: the ring-opening of 1,3-cyclohexadiene. *Phys. Chem. Chem. Phys.* 20 (26): 17714–17726.

105 Tully, J.C. (1990). Molecular dynamics with electronic transitions. *J. Chem. Phys.* 93 (2): 1061–1071.

106 Tully, J.C. (1998). Mixed quantum – classical dynamics. *Faraday Discuss.* 110: 407–419.

107 Vacher, M., Bearpark, M.J., and Robb, M.A. (2014). Communication: oscillating charge migration between lone pairs persists without significant interaction with nuclear motion in the glycine and Gly-Gly-NH-CH_3 radical cations. *J. Chem. Phys.* 140 (20): 201102.

108 Vacher, M., Mendive-Tapia, D., Bearpark, M.J., and Robb, M.A. (2014). The second-order Ehrenfest method. *Theor. Chem. Acc.* 133 (7): 1505.

109 Vacher, M., Meisner, J., Mendive-Tapia, D., Bearpark, M.J., and Robb, M.A. (2015). Electronic control of initial nuclear dynamics adjacent to a conical intersection. *J. Phys. Chem. A.* 119 (21): 5165–5172.

110 Vacher, M., Mendive-Tapia, D., Bearpark, M.J., and Robb, M.A. (2015). Electron dynamics upon ionization: control of the timescale through chemical substitution and effect of nuclear motion. *J. Chem. Phys.* 142 (9): 094105.

111 Vacher, M., Steinberg, L., Jenkins, A.J. et al. (2015). Electron dynamics following photoionization: decoherence due to the nuclear-wave-packet width. *Phys. Rev. A* 92: 040502(R).

112 Vacher, M., Albertani, F.E.A., Jenkins, A.J. et al. (2016). Electron and nuclear dynamics following ionisation of modified bismethylene-adamantane. *Faraday Discuss.* 194, 95–115.

113 Vacher, M., Bearpark, M.J., and Robb, M.A. (2016). Direct methods for non-adiabatic dynamics: connecting the single-set variational multi-configuration Gaussian (VMCG) and Ehrenfest perspectives. *Theor. Chem. Acc.* 135 (187): 1–11.

114 Vacher, M., Bearpark, M.J., Robb, M.A., and Malhado, J.P. (2017). Electron dynamics upon ionization of polyatomic molecules: coupling to quantum nuclear motion and decoherence. *Phys. Rev. Lett.* 118: 083001.

115 Werner, H.-J., Knowles, P. J., Knizia, G. et al. Molpro, version 2012.1, a package of ab initio programs.

116 Wolf, T.J.A., Sanchez, D.M., Yang, J. et al. (2019). The photochemical ring-opening of 1,3-cyclohexadiene by ultrafast electron diffraction. *Nat. Chem.* 11: 504–509.

117 Wollenhaupt, M., Engel, V., and Baumert, T. (2005). Femtosecond laser photoelectron spectroscopy on atoms and small molecules: prototype studies in quantum control. *Annu. Rev. Phys. Chem.* 56 (1): 25–56.

118 Worth, G.A. and Cederbaum, L.S. (2004). Beyond born-oppenheimer: molecular dynamics through a conical intersection. *Annu. Rev. Phys. Chem.* 55: 127.

119 Worth, G.A., Robb, M.A., and Lasorne, B. (2008). Solving the time-dependent Schrödinger equation for nuclear motion in one step: direct dynamics of non-adiabatic systems. *Mol. Phys.* 106 (16-18): 2077–2091.

120 Worth, G.A. and Burghardt, I. (2003). Full quantum mechanical molecular dynamics using Gaussian wave packets. *Chem. Phys. Lett.* 368 (3-4): 502–508.

121 Yang, J., Guehr, M., Shen, X. et al. (2016). Diffractive imaging of coherent nuclear motion in isolated molecules. *Phys. Rev. Lett.* 117: 153002.

122 Yang, J., Guehr, M., Vecchione, T. et al. (2016). Diffractive imaging of a rotational wave packet in nitrogen molecules with femtosecond megaelectronvolt electron pulses. *Nat. Commun.* 7: 11232.

123 Yonehara, T., Hanasaki, K., and Takatsuka, K. (2012). Fundamental approaches to nonadiabaticity: toward a chemical theory beyond the born–oppenheimer paradigm. *Chem. Rev.* 112 (1): 499–542.

124 Yong, H., Zotev, N., Stankus, B. et al. (2018). Determining orientations of optical transition dipole moments using ultrafast X-ray scattering. *J. Phys. Chem. Lett.* 9: 6556–6562.

125 Zewail, A.H. (2000). Femtochemistry: atomic-scale dynamics of the chemical bond. *J. Phys. Chem. A* 104 (24): 5660–5694.

126 Zewail, A.H., de Schryver, F.C., De Feyter, S., and Schweitzer, G. (2001). *Femtochemistry: with the Nobel Lecture of A. Zewail*, 1e. Wiley-VCH. ISBN: 9783527302598.

127 Zhu, C., Nangia, S., Jasper, A.W., and Truhlar, D.G. (2004). Coherent switching with decay of mixing: an improved treatment of electronic coherence for non-born–oppenheimer trajectories. *J. Chem. Phys.* 121 (16): 7658–7670.

16

Surface Hopping Molecular Dynamics

Sebastian Mai, Philipp Marquetand, and Leticia González

Institute of Theoretical Chemistry, Faculty of Chemistry, University of Vienna, Währinger Straße 17, 1090 Vienna, Austria

Abstract

In this chapter we review the wide-spread trajectory surface hopping method for non-adiabatic dynamics simulations. The surface hopping method belongs to the class of mixed quantum-classical methods that propagate electrons quantum mechanically and nuclei classically. The main focus of the review is to present the most important concepts of the surface hopping method in an accessible and practically oriented way. The covered concepts include nuclear motion, wave function propagation, decoherence correction, the actual algorithm determining the surface hops, nuclear momentum adjustment, and inclusion of diverse coupling terms. Furthermore, the chapter discusses practical aspects of surface hopping simulations, like the choice of electronic structure method, initial condition generation, and ensemble analysis. A fully worked example completes the overview over the surface hopping method.

16.1 Introduction

The description of molecular motion requires approximations in almost all cases, due to the complexity of multi-dimensional molecular wave functions. One very popular and effective approximation is the use of classical mechanics instead of quantum mechanics. The use of classical mechanics is motivated here by its favorable scaling with system size, which is in stark contrast to the exponential scaling of quantum mechanics. The classical approximation is sensible as long as quantum-mechanical effects like tunneling or interference are negligible and the energetic spacing between the quantum levels of that particle is sufficiently small compared to the kinetic energy. Since nuclei are much heavier than electrons, this spacing between quantum levels is usually much smaller for nuclear degrees of freedom than for electronic ones. Hence, it makes sense to treat the nuclei with classical mechanics while the electrons are treated with quantum mechanics. This combination of classical and quantum mechanical descriptions could be called *semiclassical* dynamics, but more common terms are *mixed quantum-classical* dynamics [1] or *ab initio* molecular dynamics, where the latter gives credit to the ab initio electronic structure calculations involved.

In general, in mixed quantum-classical dynamics, the nuclear motion is described with Newton's equation of motion $a = F/m$, i.e., acceleration equals force divided by mass [2]. The forces acting on the nuclei are determined by the attractive and repulsive interactions among the nuclei and electrons. Within the Born–Oppenheimer approximation, the forces on the nuclei are obtained as

Quantum Chemistry and Dynamics of Excited States: Methods and Applications,
First Edition. Edited by Leticia González and Roland Lindh.

the negative derivative of the electronic energy (the energy gradient) with respect to the nuclear coordinates, i.e., $F = -\frac{\partial E^{\mathrm{el}}}{\partial \vec{R}}$. Here, it can be seen that E^{el} is a function of the nuclear coordinates $\vec{R}$, which is known as a potential energy surface (PES). As long as only a single PES is considered, the resulting method is often termed *Born–Oppenheimer molecular dynamics* (note that this is a type of ab initio molecular dynamics, but often not considered to be a mixed quantum-classical method).

The situation becomes much more complicated when more than one electronic PES is involved in the physical process under investigation. This is usually the case for excited-state dynamics that occur after a molecule is electronically excited by absorption of a photon. In particular, whenever several electronic states have similar energies, the Born–Oppenheimer approximation becomes invalid and one has to include a description of non-adiabatic processes. These non-adiabatic processes occur in radiationless photo-processes such as internal conversion and intersystem crossing, but can also be present in photo-reactions involving dissociation or electron transfer. In these cases, it is *a priori* unclear how the mixed quantum-classical scheme can be applied to the problem. Several approaches have been devised for this purpose in the last decades, of which some are described in other chapters of this book.

Here, we shall present the *surface hopping* (SH) method, which is a popular approach to mixed quantum-classical dynamics including multiple electronic states. There are many excellent reviews on the topic, for example in Refs. [3–11]. Some of these focus primarily on the concepts [3, 7], some are more practically oriented [4–6], and others target new developments in the field [8–11]. In this chapter, we aim at a mixture of a conceptual and a practical review, with a focus on all important aspects necessary for beginners to carry out SH simulations on their own.

16.2 Basics of Surface Hopping

We start by giving an overview over the general ideas behind SH. This method was first proposed by Tully and Preston in the 1970s [12]. The basic assumption is that during non-adiabatic dynamics the nuclei move adiabatically for most of the time, and only for relatively short periods of time—and in relatively small regions of the configuration space—undergo non-adiabatic transitions. Hence, they proposed that, pragmatically, one could approximate the non-adiabatic transitions by instantaneous switches—or *hops*—between adiabatic PESs.

In order to illustrate this concept, in Figure 16.1, we compare the evolution of a quantum mechanical wave packet to the evolution of classical trajectories including surface hopping. As explained in the caption, the central event is the non-adiabatic transition at the avoided crossing. The quantum wave packet is split during this event, creating two branches which then evolve differently. In order to mimic this splitting, SH requires many independent trajectories, each of which undergoes stochastic hopping events.

16.2.1 Advantages and Disadvantages

Originally, SH was proposed as an ad hoc approach to non-adiabatic dynamics, without detailed theoretical justification. Since then, there were several attempts to rigorously derive SH from quantum mechanics [13, 14]. On the one hand, these approaches start at the quantum-classical Liouville Eq. [13, 14] and assume (i) unique trajectories (i.e., at any given time t, there is only one trajectory at any phase space point $(\vec{R}, \vec{p})$, excluding recoherence) [13], (ii) large nuclear velocities, and (iii) that electronic decoherence with a rate proportional to force differences is taken into account. In particular, the last point might be unclear to new readers, but it will be

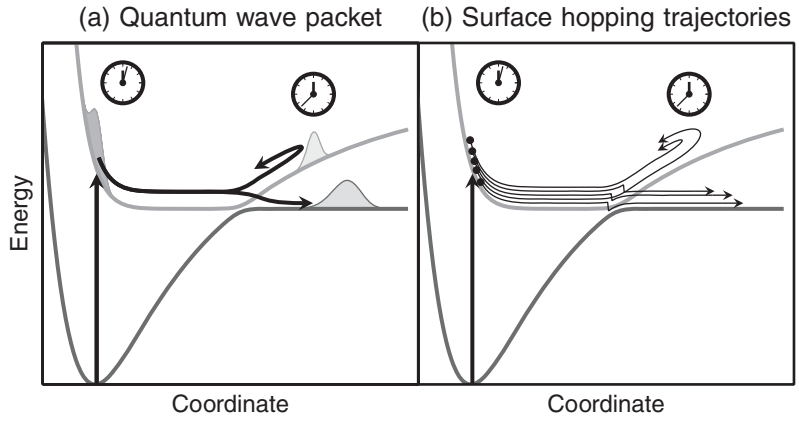


Figure 16.1 Comparison of a quantum wave packet (a) and SH trajectories (b). In (a), after the wave packet is initially promoted to the upper PES, it moves to the right, where it splits into two parts at the avoided crossing. The wave packet that crossed to the lower state continues motion to the right, whereas the wave packet that stayed on the upper state is eventually reflected by the steep PES. In (b), five SH trajectories are initially promoted to the upper state. At the avoided crossing, three trajectories hop and follow the lower PES, while two trajectories do not hop and get subsequently reflected on the upper PES.

discussed in detail below in section 16.3.3. On the other hand, the exact factorization ansatz for the molecular wave function [15, 16] (see Chapter 17) can be reduced to SH employing (i) a trajectory ansatz and (ii) neglecting quantum forces arising from coupled trajectories [10]. Besides these approximations, limitations of SH naturally arise from the classical description of nuclei, meaning that SH misses quantum effects, such as a correct description of the zero-point energy, tunneling, or nuclear interferences. Combined, these approximations mean that it is not possible to systematically converge to the exact quantum dynamical result with surface hopping, unlike more advanced methods like full multiple spawning (see Chapter 14) or direct-dynamics variational multi-configurational Gaussians (see Chapter 13).

Even though SH is not fully theoretically justified, it is among the most extensively used approaches to perform non-adiabatic dynamics simulations. Its advantages [5] include simplicity, practicality, and scalability. The first of these refers to the fact that the concepts of a classically evolving molecule and of hopping between states are intuitively accessible concepts for most chemists, which helps with the interpretation of results. The two other advantages refer to the ability of SH to deal with large polyatomic systems efficiently and to the possibility of computing many independent trajectories in parallel. Finally, it has been shown by several groups that SH (including decoherence) can provide accurate results compared to exact methods for a large number of model systems [17–19], provided an adequate level of theory is employed for the electronic structure calculations.

16.2.2 General Algorithm

Within SH, the nuclei follow Newton's classical equation of motion:

$$M_A \frac{\partial^2 \vec{R}_A}{\partial t^2} = -\frac{\partial E^{\text{el}}}{\partial \vec{R}_A}, \tag{16.1}$$

where M_A and $\vec{R}_A$ are the mass and position of nucleus A, and E^{el} is the current electronic energy. In practice, the nuclear velocity $\vec{v}_A$ is also explicitly included. The temporal evolution of the position of all nuclei $\vec{R}(t)$ is then referred to as a *trajectory*.

Besides the nuclei, in SH an electronic wave function is also considered. This wave function is expressed as a linear combination of basis states:

$$|\Phi^{\text{el}}(t)\rangle = \sum_{\alpha} c_{\alpha}(t)|\Psi_{\alpha}(t)\rangle, \tag{16.2}$$

where α runs over the basis states, $c_{\alpha}(t)$ are time-dependent coefficients, and $|\Psi_{\alpha}(t)\rangle$ are the basis states. Most often, these basis states are adiabatic states, but this topic will be discussed in more detail in section 16.3.6. Now it is only important to realize that the temporal evolution of this wave function follows the time-dependent Schrödinger equation and that it is affected by the $\vec{R}(t)$ through the parametric dependence of the electronic Hamiltonian on $\vec{R}(t)$. In turn, the evolution of $\vec{R}(t)$ depends on the gradient of the electronic energy E^{el} and thus on the evolution of $|\Phi^{\text{el}}(t)\rangle$, showing that the classical nuclear evolution and the quantum-mechanical electronic evolution are intimately coupled.

The specific connection between the electronic wave function and the electronic gradient is the defining aspect of SH. In SH, the electronic gradient is given by the gradient of the *active* state. The active state, denoted by β in the following, is chosen stochastically in each time step such that across the trajectory ensemble the fraction of trajectories with active state β is equal to the electronic population of that state:

$$\frac{N_{\beta}(t)}{N_{\text{traj}}} = \frac{1}{N_{\text{traj}}} \sum_{i}^{N_{\text{traj}}} |c_{\beta}^{i}(t)|^2, \tag{16.3}$$

where i runs over all trajectories. Accordingly, if the electronic population $|c_{\beta}(t)|^2$ changes, a corresponding number of trajectories will hop from/to β to fulfill Eq. (16.3). More details on the actual hopping algorithm will be given below.

In a very simplistic way the SH algorithm for one time step is composed of the following computations:

1. Calculate the new positions of the nuclei,
2. Compute electronic quantities: energies, gradients, couplings,
3. Calculate the new electronic coefficients,
4. Choose the new active state and obtain the corresponding forces,
5. Continue at step 1.

A much more detailed tutorial for an SH algorithm is given in the next section, where special attention is also paid to the first time step. Here, we finally want to draw attention to step 2, where the electronic quantities are computed. There are two general approaches to obtain these. Historically, before an SH simulation was performed, the full form of the PES was defined, usually through parametrized functions. All electronic properties were then computed directly from these functions. However, nowadays most SH simulations are performed *on-the-fly*, meaning that no pre-defined PESs are used but all required quantities are computed for the current time step, using semi-empirical, density-functional-based, or ab initio electronic structure methods. The on-the-fly scheme has been pivotal in the success of SH, as it eliminates the non-trivial and very time-consuming step of preparing an accurate and complete PES, as it is necessary for propagating quantum wave packet dynamics (see Chapters 11 and 12).

16.3 Surface Hopping Ingredients

While the basic idea of SH—as explained above—is very simple, recent decades have seen dozens of variants of SH. In general, these variants modify some part of the SH algorithm, e.g., the equation used to compute the hopping probabilities or the way in which the total energy is conserved during a hop. These modifications are usually applied in order to improve the agreement with quantum dynamics reference computations, or in order to fulfill some theoretical requirement. Hence, in this section, we describe in detail all components of the surface hopping method and the most common techniques applied.

For convenience, in Figure 16.2 we provide an overview of the most important quantities that are involved in the SH simulations, how they evolve from time step t to time step $t + \Delta t$, and in which order these computations can be carried out. These steps are (1) update the nuclear positions $\vec{R}$, (2) compute the gradients $\vec{\nabla} E_\beta$, energies (in matrix $\mathbf{H}$), and couplings (matrix $\mathbf{K}$), (3) update the nuclear velocities $\vec{v}$ with the newly computed forces via the acceleration $\vec{a}$, (4) propagate the electronic coefficients $\vec{c}$, (5) compute the hopping probabilities $\vec{h}$, (6) stochastically select the new active state β, and (7) compute the gradient of the new active state for the next iteration. In the following subsections, some of these steps are explained in more detail.

Naturally, in order to start the iterations shown in Figure 16.2 it is necessary to obtain all quantities for $t = 0$. This involves two steps. First, the preparation of initial conditions according to the physical or chemical problem at hand provides values for $\vec{R}(0)$, $\vec{v}(0)$, $\vec{c}(0)$, and $\beta(0)$. Second, an electronic structure calculation provides values for $\mathbf{H}(0)$, $\mathbf{K}(0)$, and $\vec{\nabla} E_\beta(0)$. These two steps are discussed further below in section 16.4.

16.3.1 Nuclear Motion

In SH, one of the most ubiquitous algorithms to solve Newton's equation is the velocity-Verlet algorithm [20]. In this algorithm, first the nuclear positions are updated with the following equation:

$$\vec{R}_A(t + \Delta t) = \vec{R}_A(t) + \vec{v}_A(t)\Delta t + \frac{1}{2}\frac{\nabla_A E_\beta(t)}{M_A}\Delta t^2. \tag{16.4}$$

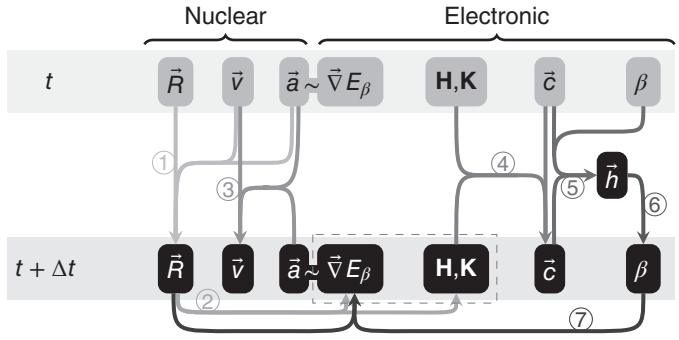


Figure 16.2 Overview of the computations carried out during one step (from t to $t + \Delta t$) of a SH simulation. The seven steps are described in the text. The dashed box denotes the results of the electronic structure calculation. Not shown here are the modification of $\vec{c}$ due to the decoherence correction scheme and the modification of $\vec{v}$ due to momentum adjustments after a hop.

Note how the right-hand side of this equation only depends on quantities from step t, which means that this equation is naturally the very first one carried out at the beginning of a time-step iteration. In Figure 16.2, it is denoted as (1).

The second half of the velocity-Verlet algorithm is the update of the nuclear velocities:

$$\vec{v}_A(t+\Delta t) = \vec{v}_A(t) + \frac{1}{2}\frac{\vec{\nabla}_A E_\beta(t) + \vec{\nabla}_A E_\beta(t+\Delta t)}{M_A}\Delta t. \tag{16.5}$$

This step is denoted as (3) in Figure 16.2. Note how here the gradient $\vec{\nabla}_A E_\beta(t+\Delta t)$ is required, which means that the electronic structure calculations (2) have to be performed between (1) and (3). For proper energy conservation, in both equations the same state β should be considered. This means that the decision to hop is taken after the velocities have been updated. The gradient of the new active state hence only enters the propagation in the next iteration. Hence, this step is denoted in Figure 16.2 as (7).

16.3.2 Wave Function Propagation

By inserting the electronic wave function expansion into the time-dependent Schrödinger equation, the equation of motion for the electrons can be obtained:

$$\frac{\partial}{\partial t}c_\beta = -\sum_\alpha \left[\frac{\mathrm{i}}{\hbar}H_{\beta\alpha} + K_{\beta\alpha}\right] c_\alpha, \tag{16.6}$$

or in matrix notation:

$$\frac{\partial}{\partial t}\vec{c} = -\left[\frac{\mathrm{i}}{\hbar}\mathbf{H} + \mathbf{K}\right]\vec{c}, \tag{16.7}$$

where $\hbar$ is the reduced Planck constant, $H_{\alpha\beta} = \langle\Psi_\alpha|\hat{H}|\Psi_\beta\rangle$ and $K_{\alpha\beta} = \langle\Psi_\alpha|\partial/\partial t|\Psi_\beta\rangle = \partial\vec{R}/\partial t \cdot \langle\Psi_\alpha|\partial/\partial\vec{R}|\Psi_\beta\rangle = \vec{v}\cdot\langle\Psi_\alpha|\vec{\nabla}|\Psi_\beta\rangle$. Given $\vec{c}(t)$, $\mathbf{H}(t)$, $\mathbf{H}(t+\Delta t)$, $\mathbf{K}(t)$, and $\mathbf{K}(t+\Delta t)$, this equation can be integrated to find $\vec{c}(t+\Delta t)$. In Figure 16.2, this step is denoted as (4). The numerical integration can be accomplished, e.g., with Runge–Kutta/Butcher algorithms or with short-time matrix exponential methods, as described for example in [4]. The short-time matrix exponential method can be written as:

$$\vec{c}(t+\Delta t) = \underbrace{\hat{\mathcal{T}}\exp\left[-\int_t^{t+\Delta t}\left(\frac{\mathrm{i}}{\hbar}\mathbf{H}(\tau) + \mathbf{K}(\tau)\right)\mathrm{d}\tau\right]}_{\mathbf{P}(t+\Delta t,t)}\vec{c}(t). \tag{16.8}$$

where $\hat{\mathcal{T}}$ is the time-ordering operator and $\mathbf{P}(t+\Delta t)$ is the propagator (time evolution) matrix.

While the above equation of motion arises naturally from the Schrödinger equation, it is numerically difficult to integrate, mainly because $\mathbf{K}(t)$ can change extremely quickly locally. Furthermore, $\mathbf{K}(t)$ is usually computed from the non-adiabatic coupling (NAC) vectors $\langle\Psi_\alpha|\vec{\nabla}|\Psi_\beta\rangle$, which are often not available in current implementations of several electronic structure methods (e.g., most single-reference methods). Hence, many SH implementations compute $\mathbf{K}(t)$ approximately from the numerical differentiation of wave function overlaps [21, 22]:

$$K_{\alpha\beta}\left(t+\frac{\Delta t}{2}\right) \approx \frac{1}{2\Delta t}(S_{\alpha\beta}(t,t+\Delta t) - S_{\alpha\beta}(t+\Delta t,t)). \tag{16.9}$$

with the overlaps $S_{\alpha\beta}$ defined as:

$$S_{\alpha\beta}(t,t+\Delta t) = \langle\Psi_\alpha(t)|\Psi_\beta(t+\Delta t)\rangle. \tag{16.10}$$

The advantage of this scheme is that the overlaps can always be computed if a configuration interaction-like representation of the electronic states is known [23–25].

An even more sophisticated method is the so-called *local diabatization* scheme [27, 28], which has been shown [28] to be numerically more stable than the normal scheme of integrating **K**. This stability is achieved by avoiding the combination of numerical differentiation of $\mathbf{S}(t, t+\Delta t)$ (Eq. (16.9)) to obtain **K** and subsequent integration of this **K** (Eq. (16.8)). These numerical problems with the normal scheme of integrating **K** are exemplified in Figure 16.3, where it is shown that narrow peaks in **K**, found at so-called "trivial crossings" [26], might be missed if the time step is too long. Within the local diabatization scheme, **K** or the NAC vectors are never explicitly computed. Instead, the overlap matrix **S**—which is closely related to the integral of **K**—is directly employed in propagating the wave function coefficients $\vec{c}$:

$$\vec{c}(t+\Delta t) = \underbrace{\mathbf{S}(t, t+\Delta t)^{\dagger}\hat{\mathcal{T}} \exp\left[-\int_{t}^{t+\Delta t} \frac{\mathrm{i}}{\hbar}\mathbf{S}(t,\tau)\mathbf{H}(\tau)\mathbf{S}(t,\tau)^{\dagger}\mathrm{d}\tau\right]}_{\mathbf{P}(t+\Delta t,t)}\vec{c}(t). \tag{16.11}$$

Here, in the expression $\mathbf{S}(t,\tau)\mathbf{H}(\tau)\mathbf{S}(t,\tau)^{\dagger}$ the Hamiltonian matrix **H** is diabatized, which is the origin of the name of the method.

16.3.3 Decoherence

Only a few years after the publication of the original SH algorithm [29], it was realized that there is one systematic problem in SH [30, 31]. This problem is that the electronic populations are propagated with *too much* coherence [9, 13, 29, 32, 33].

An example of this problem is shown in Figure 16.4(a), where a single trajectory passes through a double crossing. Initially, the trajectory is moving on the upper state, with 100% of the electronic amplitudes corresponding to that state. After the first crossing, 60% of the amplitude is transferred to the lower state, inducing a surface hop. From that point, the trajectory follows the gradient of the lower state, and hence the trajectory is able to move further to the right. In contrast, if no hop had occurred, the trajectory would have been reflected at the classically forbidden barrier in the upper state. The central problem of the shown trajectory is that the amplitudes on the upper state

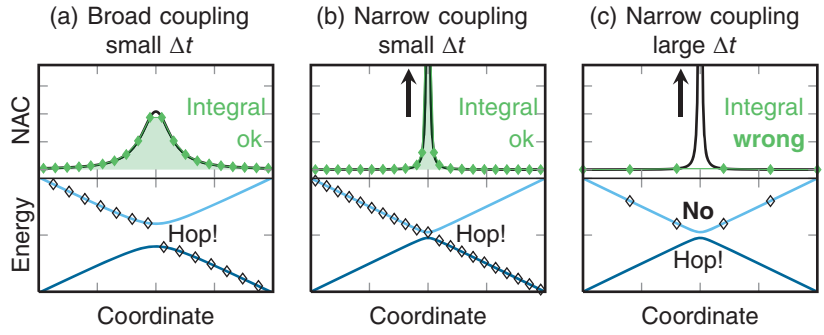


Figure 16.3 Influence of the width of the non-adiabatic coupling and the time step on SH. If the coupling is broad (a), then it can be integrated accurately and the non-zero integral might induce a hop. However, if the coupling is narrow (b and c), then short time steps Δt are needed for an accurate integration (b). Otherwise, the narrow peak is missed, leading to a too small value of the integral and failure to perform the hop (c). In the literature, this problem of strong, narrow couplings is sometimes called the "trivial crossing problem" [26].

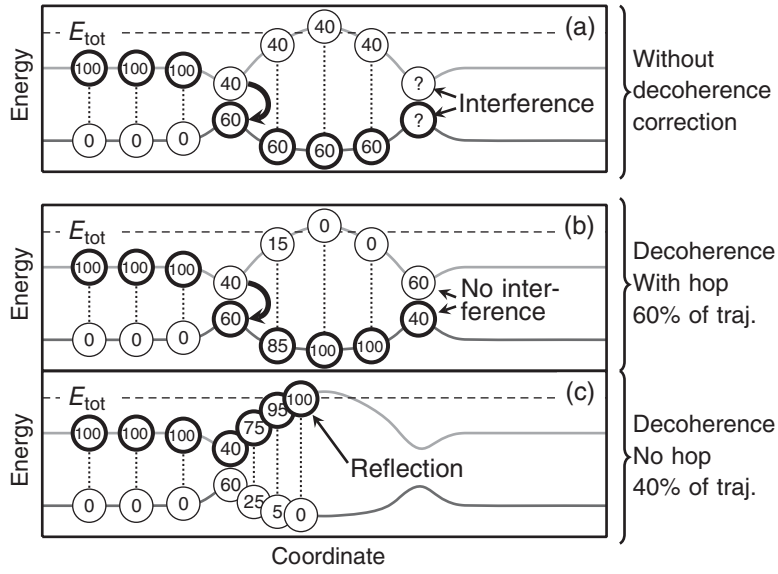


Figure 16.4 Sketch explaining the problem of decoherence in SH. It shows a single trajectory passing through a double crossing. The numbers in the circles denote the population (in %) in the respective state at each time step, with a thick circle denoting the active state. Without decoherence correction (a), the population in the upper state is wrongly dragged over the barrier because the lower state is the active state. When lower and upper states interact at the second crossing, artificial interference arises. In (b) and (c), the problem is circumvented by applying a decoherence correction, see text.

(40%) also follow the gradient of the lower state and move to the right, although this should be classically forbidden. Once the trajectory arrives at the second crossing point, there will be electronic interference between the 40% on the upper state and the 60% on the lower one. This interference will strongly affect the amount of transferred population depending on the relative complex phase between the two states. However, as the upper amplitudes should not have reached this crossing at all, no interference should occur.

One way to solve this problem—electronic amplitudes following the gradient of another state—is to split the trajectories when a hop occurs. This is basically the idea that led to the development of the multiple spawning [34] and multiple cloning [35] methods (see Chapters 14 and 15), which are more accurate alternatives to SH because they treat the nuclei quantum-mechanically. When in the SH framework, decoherence can instead be realized by a decoherence correction algorithm that over time collapses or decreases the amplitudes of the non-active states, so that far away from crossing regions, every trajectory eventually recovers a pure wave function. This is shown in Figure 16.4(b) and c, where non-active amplitudes are damped. In Figure 16.4(b), after the hop to the lower state, the decoherence correction destroys the population of the upper state, so that no interference occurs at the second crossing. It is important here to note that in our example the total population across all trajectories (e.g., N_{traj}=100) does not change: at the first crossing, there are 100 trajectories with 40% in the upper state and 60% in the lower one. After the decoherence correction has taken full effect, there are 60 trajectories with 100% in the lower state. Additionally, as shown in Figure 16.4(c), there are 40 trajectories with 100% in the upper state, showing that the total electronic populations did not change.

In practice, different decoherence schemes exist for SH which differ in the way the decay of the non-active amplitudes is modeled. Early approaches [36] simply set the amplitude of a non-active

state to zero as soon as the coupling with the active state falls below a predefined threshold [37]. A more elaborate and very popular approach for decoherence is based on the "decay of mixing" concept [32, 38] and its descendents [33, 39]. There, the non-active amplitudes are damped exponentially in time, with the decay rate dependent on the kinetic energy—such that no decoherence occurs if the nuclei are not moving—and the energy gap to the active state, with larger gaps inducing faster decay. No additional quantities are required, making this approach very simple and explaining its widespread use [11, 40, 41].

The most advanced decoherence correction schemes introduce auxiliary trajectories for each electronic state [33, 42–44]. These extra variables propagate along each principal trajectory and represent how fast the active and non-active amplitudes separate in phase space, thus allowing a realistic decoherence rate to be estimated. The big advantage of some of these schemes [42–44] is that the decoherence rate is proportional to the difference in gradients between the different states, which is the theoretically correct relationship [9, 30, 31]. This was shown by the above-mentioned derivations of the surface hopping algorithm from first principles either through the quantum-classical Liouville equation or through the exact factorization method [13, 14, 16].

While the decoherence correction might seem like a rather technical detail of SH, it has a quite tangible effect—for an ensemble of trajectories, it enforces the "internal consistency" that is described by Eq. (16.3). In other words, it guarantees that the fraction of trajectories in each state (called the "classical" population; left-hand side of (16.3)) is consistent with the average of the squared electronic amplitudes (the "quantum" population; right-hand side of (16.3)). As shown above, decoherence corrections also make the electronic propagation much more accurate in any system where a trajectory passes through more than one crossing region. Hence, nowadays all SH simulations should employ some decoherence correction scheme, optimally one that incorporates the difference gradients [42–44].

16.3.4 Surface Hopping Algorithm

Besides the mentioned decoherence correction, SH trajectories are largely influenced by the actual algorithm that is used for the hopping procedure. There are three main aspects which need to be considered here.

First, there are two main concepts to decide *when* to hop. One approach is to hop only at specific geometries or points along the trajectory, defined either prior to the simulations (e.g., optimized crossing points) or detected during the simulation (e.g., minimum in the energy gap, maximum in the coupling term). This approach is historically the oldest, but still in use, for example in the context of the Zhu–Nakamura SH method [45–47]. However, the approach is not very extensively used nowadays, because it is often not practical to define all hopping points a priori, and hopping point detection can become complicated or ill-defined for systems with many states. Instead, many SH implementations allow surface hops in principle in any simulation time step [29]. This is especially advantageous if hops can occur in a large phase space volume, where it is not possible to determine the single most important point. Thus, in the following we will only consider SH algorithms that allow transitions to take place at any time step.

The second aspect of the SH algorithm is the equation that governs the hopping probabilities for each time step. The process of computing these probabilities is shown in the overview in Figure 16.2 as (5). Most reported equations incorporate the fewest-switches criterion [29], which demands that Eq. (16.3) has to be fulfilled with the smallest possible number of hops. This restriction was introduced after it was realized that excessive hopping in SH makes the trajectories move according to the average of several PESs, which is unphysical. Hence, as long as the electronic populations

do not change, no hop should be performed. This fewest switches criterion can be satisfied by never hopping out of state β if $|c_\beta|^2$ increases and otherwise hopping out of this state with total probability:

$$p_{\beta\rightarrow} = 1 - \frac{|c_\beta(t+\Delta t)|^2}{|c_\beta(t)|^2}. \tag{16.12}$$

For example, this equation means that if $|c_\beta(t+\Delta t)|^2$ is half as large as $|c_\beta(t)|^2$, the probability is 50% to hop out of state β.

Unfortunately, the fewest-switches criterion does not provide any guidance on how this total probability should be partitioned over the possible target states α. Hence, there are different recipes that can be used. In Tully's original formulation [29], assuming a short time step Δt, one can make the approximation:

$$p_{\beta\rightarrow} \approx -\frac{2\Delta t}{|c_\beta(t)|^2}\Re\left(c_\beta^*(t)\frac{\partial c_\beta(t)}{\partial t}\right). \tag{16.13}$$

The term $\partial c_\beta(t)/\partial t$ can then be substituted by the right-hand side of Eq. (16.6):

$$p_{\beta\rightarrow} \approx -\frac{2\Delta t}{|c_\beta(t)|^2}\Re\left(c_\beta^*(t)\sum_\alpha\left[\frac{\mathrm{i}}{\hbar}H_{\beta\alpha}+K_{\beta\alpha}\right]c_\alpha(t)\right). \tag{16.14}$$

The expressions for each individual $p_{\beta\rightarrow\alpha}$ are then simply obtained by splitting the sum over α.

Other approaches prefer to use the expression in Eq. (16.12) and partition it either based on the propagator matrix in Eq. (16.8) [27, 48], e.g.,:

$$p_{\beta\rightarrow\alpha} = \left(1 - \frac{|c_\beta(t+\Delta t)|^2}{|c_\beta(t)|^2}\right)\frac{\Re[c_\alpha(t+\Delta t)P^*_{\alpha\beta}c_\beta(t)]}{|c_\beta(t)|^2 - \Re[c_\beta(t+\Delta t)P^*_{\beta\beta}c_\beta(t)]}, \tag{16.15}$$

or based solely on increases and decreases of the squared amplitudes [49, 50].

The third aspect regards the stochastic sampling carried out based on the computed probabilities, and the selection of a new active state for the next time step. This selection is shown in Figure 16.2 as (6). Nowadays, almost all SH implementations use the "anteater" technique [12], where a random number r between 0 and 1 is drawn in each time step. A hop from β to α is performed if:

$$\sum_{i=1}^{\alpha-1} h_{\beta\rightarrow i} < r \leq h_{\beta\rightarrow\alpha} + \sum_{i=1}^{\alpha-1} h_{\beta\rightarrow i} \tag{16.16}$$

with random number r between 0 and 1. This stochastic scheme is exemplified in Figure 16.5. The name of this technique comes from the analogy with the anteater (the trajectory) that tends to follow the path where most ants (the electronic population) walk. There also exist alternative sampling schemes, for example the "army ants" algorithm [51] where the hopping decision is made

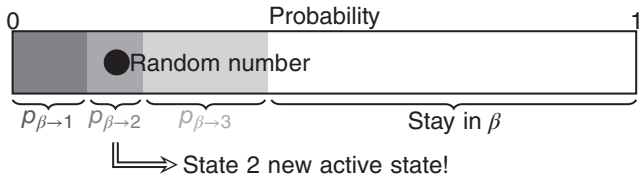


Figure 16.5 Sketch explaining the anteater technique to decide on a new state based on computed hopping probabilities. In this example, the random number falls within the interval corresponding to state 2, which thus becomes the new active state, except if the hop is frustrated (see below).

from uniform probabilities (i.e., hops to all states are equally likely) but the values $h_{\beta\to\alpha}$ are stored to weigh the trajectories during final analysis. This technique might be useful for better sampling of rare reaction channels if they are the focus of the investigation.

To end this subsection, some more specialized variants for the hopping probability will be mentioned. An interesting example is "decoherence-induced surface hopping" [52], where the decoherence correction is incorporated in the electronic equation of motion as a dissipative term, and surface hops only occur if the current active state is collapsed by decoherence. Another special surface hopping method is Liouville-space SH [53], where in addition to the adiabatic surfaces, trajectories may hop to *coherence surfaces* if $\rho_{\alpha\beta} = c_\alpha^* c_\beta$ is large. These surfaces are computed simply as the average of the two involved states, and are typically only populated in interaction regions.

In contrast to these sophisticated hopping schemes, an approach that is sometimes used [10, 54–56] is to assume a unit hopping probability if the energy gap between two states becomes small. The latter approach is sometimes used if the electronic structure method is not able to describe crossings between S_0 and S_1, like most single-reference methods. However, much caution has to be applied with this severe approximation, as the actual dynamics close to the intersection seam is not simulated and the decay rates might be significantly overestimated [56].

16.3.5 Kinetic Energy Adjustment and Frustrated Hops

After the decision has been made to hop to a new active state, it is usually necessary to adjust the kinetic energy of the nuclei. This step is required because the new active state will have a different potential energy than the previous active state, but the total energy (i.e., kinetic plus potential energy) needs to be conserved. The kinetic energy is adjusted by modifying the velocity vector $\vec{v}$ such that the new kinetic plus potential energy equals the total energy. This is sketched in Figure 16.6(a) and (b) for the cases of increasing or decreasing kinetic energy.

The simplest prescription for the kinetic energy adjustment is to rescale the complete velocity vector $\vec{v}$ according to:

$$\vec{v}^{\text{ adjusted}} = \sqrt{\frac{E_{\text{total}} - E_\beta}{E_{\text{total}} - E_\alpha}}\vec{v}, \tag{16.17}$$

where E_α is the potential energy of the old active state, E_β the potential of the new active state, and E_{total} the total energy. As can be seen, no additional quantities (like NAC vectors) are required. This approach is, therefore, regularly employed [23, 57, 58], but one should keep in mind that it makes the SH algorithm size-inconsistent. By size-inconsistent we mean the following: imagine two SH simulations, the first for a molecule A and the second for the same molecule A plus another molecule B at infinite distance. In both simulations, the dynamics of molecule A should be identical, since A and B are not interacting. At some time step, in A the decision is made to hop to a state with an energy that is higher than the total energy of A. Naturally, in the first system (only A present) the hop is frustrated and cannot be made. However, in the second system ($A + B$), the total energy of the system is large enough to permit the hop. In this situation, the rescaling of the full velocity vector is unphysical, because the energy to hop between states of A is paid by slowing down the atoms of B, although the two systems should not interact. This general problem does not only occur for infinitely-separated systems, but also for finitely-separated systems containing weakly interacting subsystems. The problem is especially severe in large systems like chromophores in solution or bio-environments. Here, an ad hoc solution is to only apply the velocity adjustment to atoms of the chromophore while ignoring the kinetic energy of the environment atoms [11].

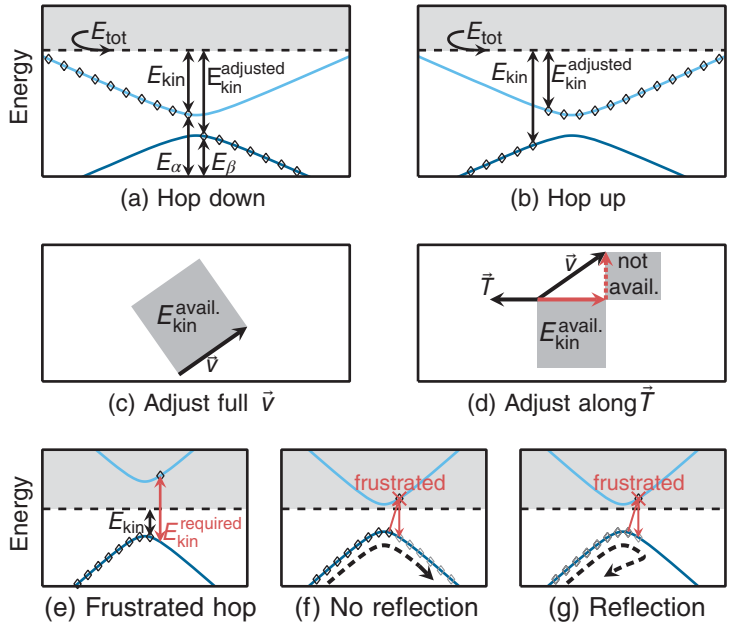


Figure 16.6 Sketches of different scenarios that can occur during kinetic energy adjustment. The top row shows simple examples of hops where the potential energy of the active state (a) decreases during the hop or (b) increases. In the middle row, the available kinetic energy for (c) full $\vec{v}$ scaling is compared to (d) adjustment only parallel to the NAC vector $\vec{T}$. As can be seen, in the latter less energy is available to make the hop. In the bottom row, we show an example of a frustrated hop (e), where the total energy is lower than the potential energy of the designated new active state. When a frustrated hop is detected, the active state is not changed, and the velocity vector is either unchanged (f) or the component along $\vec{T}$ is reflected.

The instantaneous modification of the velocity vector can be regarded as the application of an infinitely large force to the nuclei. The correct direction of that force vector can be shown—by invoking the effective quantum force method of Pechukas [59, 60]—to be the direction of the NAC vector that induced the transition from the old state to the new state. This approach is rigorous [61–63] and size-consistent. The equation for rescaling $\vec{v}$ is much more involved than the one above and can be found in the literature, e.g., Ref. [4]. Its biggest disadvantage is that it requires the computation of the relevant NAC vector, which might be computationally expensive or not even available in all electronic structure programs.

In any case, during the velocity rescaling there is only a certain amount of energy available to "pay" for hops to states with higher potential energy. When rescaling is applied to the full $\vec{v}$, the full nuclear kinetic energy is available and hops are possible to all states whose potential energy is below the total energy. However, if rescaling is applied parallel to the NAC vector, only kinetic energy along this vector is available for hops. This difference is exemplified in Figure 16.6(c) and (d). If this available energy is less than the energy difference between the old active state and the designated new active state picked by the hopping procedure, then the hop cannot be performed. Such a hop is called "frustrated", and is rejected during the SH algorithm, so that the trajectory simply continues on the previous active state without change. An interesting alternative approach here is "surface hopping with time uncertainty" [64], where trajectories are allowed to hop slightly earlier or later (according to the energy-time uncertainty $\Delta E \Delta t \approx \hbar$) if the hop is forbidden at the current time step. Some authors also have suggested that after a hop is rejected, the component of

$\vec{v}$ in the direction of the relevant NAC vector should be inverted [21, 65, 66], whereas other authors report that this reflection should not be carried out [57]. The concepts of frustrated hops and of no reflection/reflection are visualized in Figure 16.6(e–g).

16.3.6 Coupling Terms and Representations

In the electronic equation of motion (16.7), the Hamiltonian and non-adiabatic coupling matrix elements appear. These can be written as

$$H_{\beta\alpha}(t) = \langle \Psi_\beta(t) | \hat{H} | \Psi_\alpha(t) \rangle \tag{16.18}$$

and

$$K_{\beta\alpha}(t) = \sum_A \vec{v}_A \cdot \langle \Psi_\beta(t) | \vec{\nabla}_A | \Psi_\alpha(t) \rangle, \tag{16.19}$$

respectively. These quantities are affected by (i) the choice of the electronic Hamiltonian operator $\hat{H}$ and (ii) the choice of the basis functions $\Psi_\alpha(t)$ and $\Psi_\beta(t)$. The choice of the electronic Hamiltonian fundamentally determines which processes can be described with the SH simulation. The simplest electronic Hamiltonian is the molecular Coulomb Hamiltonian (MCH), which is written as:

$$\begin{aligned} \hat{H}^{\mathrm{MCH}} = &- \sum_i \frac{\hbar^2}{2m_\mathrm{e}} \nabla_i^2 \\ &+ \frac{e^2}{4\pi\varepsilon_0} \left[\sum_{A<B} \frac{Z_A Z_B}{|\vec{R}_A - \vec{R}_B|} - \sum_A \sum_i \frac{Z_A}{|\vec{R}_A - \vec{r}_i|} + \sum_{i<j} \frac{1}{|\vec{r}_i - \vec{r}_j|} \right], \end{aligned} \tag{16.20}$$

where e is the elementary charge, $\frac{1}{4\pi\varepsilon_0}$ is the Coulomb constant, Z are the nuclear charges, A and B run over nuclei, and i and j run over electrons. Using this Hamiltonian, it is possible to describe adiabatic and internal conversion dynamics for molecules. To include further processes, one needs a Hamiltonian which contains additional terms:

$$\hat{H}^{\mathrm{full}} = \hat{H}^{\mathrm{MCH}} + \hat{H}^{\mathrm{additional}}. \tag{16.21}$$

For example, for ISC $\hat{H}^{\mathrm{additional}}$ should include a spin–orbit operator, like the Breit–Pauli operator [67]:

$$\begin{aligned} \hat{H}^{\mathrm{SO,BP}} = &\frac{\hbar e^2}{2m_\mathrm{e}^2 c^2} \left[\sum_i^{n_\mathrm{el}} \sum_A^{n_\mathrm{nuc}} \frac{Z_A (\vec{r}_{iA} \times \vec{p}_i) \cdot \vec{s}_i}{r_{iA}^3} \right. \\ &\left. - \sum_{i \neq j}^{n_\mathrm{el}} \frac{(\vec{r}_{ij} \times \vec{p}_i) \cdot \vec{s}_i}{r_{ij}^3} + 2 \sum_{i \neq j}^{n_\mathrm{el}} \frac{(\vec{r}_{ij} \times \vec{p}_i) \cdot \vec{s}_j}{r_{ij}^3} \right], \end{aligned} \tag{16.22}$$

where c is the speed of light, and $\vec{r}$, $\vec{p}$, and $\vec{s}$ are distance, momentum, and spin vectors, respectively. Instead of using the full Breit–Pauli operator that is very expensive to evaluate, nowadays much work is performed with mean-field one-electron approximations to the Breit–Pauli operator [68–71]. For absorption or stimulated emission, instead, one requires the interaction between an electric field and the charge distribution of the system. In the usual dipole approximation, the corresponding Hamiltonian to be included is:

$$\hat{H}^{\mathrm{DFC}} = \left(- \sum_i^{n_\mathrm{el}} \vec{r}_i + \sum_A^{n_\mathrm{nuc}} Z_A \vec{R}_A \right) \cdot \vec{\mathcal{E}}, \tag{16.23}$$

where DFC is an acronym for dipole-field coupling. Other choices of additional Hamiltonians are also possible, for example by adding also a polarization-square of field term, or a term allowing for spontaneous emission [72].

The choice of the basis functions $|\Psi_\alpha(t)\rangle$ in Eq (16.2) (this choice is usually called *representation*) also significantly affects the form of the matrix elements (16.18) and (16.19). For example, when represented in a diabatic basis, **H** will generally be non-diagonal and **K** will be zero. On the other hand, if the basis is an eigenbasis of the full Hamiltonian, then **H** will be diagonal and **K** will be non-zero. We want to note here that—unlike in nuclear quantum dynamics (see Chapters 11 and 12)—in SH the choice of the basis is not just important for an efficient and accurate numerical treatment, but actually affects the physical outcome of the simulations. There are several reasons for this. First, in SH the trajectories always follow the gradients of the PESs of the basis states, but never the gradient of a linear combination. As shown in Figure 16.7, different representations lead to different PESs. This can, e.g., mean that a reaction barrier is classically forbidden in one basis, but allowed in another basis (compare the barrier between the two low-lying minima in Figure 16.7(a/b) with (c)). Second, all decoherence correction schemes enforce that eventually the electronic wave function becomes a pure basis state, never a linear combination. Third, if spin–orbit couplings (SOCs) are considered, then the hopping probabilities into the components of multiplets depend on the representation [73]. Furthermore, in different representations the electronic couplings can be either localized or delocalized, which affects how many trajectories are required to properly sample all reaction pathways.

For all these reasons, the choice of the optimal representation for SH is a topic discussed at length in the literature [1, 9, 14, 48, 73–75]. Generally, the best representation seems to be the adiabatic one, i.e., the one that diagonalizes the full electronic Hamiltonian including all additional couplings. Unfortunately, nowadays electronic structure codes cannot routinely compute excited eigenstates of the full electronic Hamiltonian including SOC or other additional couplings, especially for multiple excited states including gradients and non-adiabatic couplings, with few exceptions [76, 77]. Hence, some SH codes [11, 78]–including the SHARC (surface hopping including arbitrary couplings) package [11]–resort to a two-step approach that is based on quasi-degenerate perturbation theory [79]. In the first step, a desired number of MCH eigenstates $\{\Psi_\alpha^{\mathrm{MCH}}\}$ are found (e.g., the few lowest singlet and triplet states). For these MCH states, electronic structure codes can routinely deliver gradients, dipole moments, and all relevant couplings. In the second step, these quantities are transformed from the MCH basis to the diagonal one. To this end, the matrix representation of

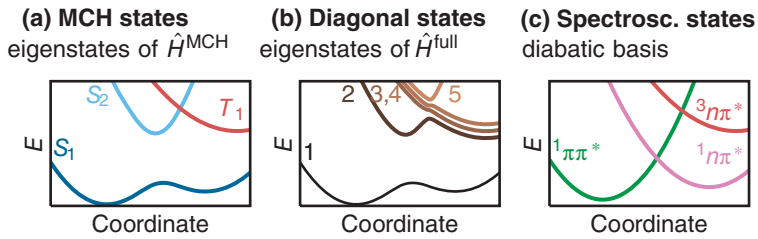


Figure 16.7 Example PESs for the three most common representations used in SH. The MCH states (a) are usually the ones computed with electronic structure methods. The diagonal states (b) are obtained by specialized electronic structure or by a posteriori diagonalization. The spectroscopic states (c) consist of coordinate-independent basis states, which generally do not exist for poly-atomic molecules. Adapted from Ref. [11]: Mai, Marquetand, González, *WIREs: Comp. Mol. Sci.* 8 e1370 (2018) under CC-BY license.

the full Hamiltonian in the basis of these MCH states is constructed and diagonalized:

$$\underbrace{\langle \Psi_\alpha^{\text{diag}} | \hat{H}^{\text{full}} | \Psi_\beta^{\text{diag}} \rangle}_{\delta_{\alpha\beta} H_{\alpha\beta}^{\text{diag}}} = \sum_\mu \sum_\nu \underbrace{\langle \Psi_\alpha^{\text{diag}} | \Psi_\mu^{\text{MCH}} \rangle}_{U_{\mu\alpha}^\dagger} \underbrace{\langle \Psi_\mu^{\text{MCH}} | \hat{H}^{\text{full}} | \Psi_\nu^{\text{MCH}} \rangle}_{H_{\mu\nu}^{\text{MCH}}} \underbrace{\langle \Psi_\nu^{\text{MCH}} | \Psi_\beta^{\text{diag}} \rangle}_{U_{\nu\beta}}, \tag{16.24}$$

where the $H_{\mu\nu}^{\text{MCH}}$ elements are computed by the electronic structure code after the $\{\Psi_\alpha^{\text{MCH}}\}$ have been found. The transformation matrix $\mathbf{U}$ with elements $U_{\alpha\beta}$ can then be used to transform all required properties like the gradients and couplings. The approximation here is that the diagonalization step only includes the small number of considered states instead of all possible states. In many cases, this is a good approximation, especially when considering SOC without very heavy atoms [80, 81].

16.4 Practical Remarks

Besides the actual SH algorithm and the different possible ingredients reviewed above, the success of an SH study critically depends on a number of further aspects. Two very important issues need to be considered before the SH simulations can even be started. One is the choice and validation of the electronic structure level of theory for the on-the-fly calculations. After a suitable level of theory has been found, the next step is to prepare suitable initial conditions, which are typically chosen to represent some (hypothetical) experimental setup. Only once these initial conditions are ready, can the dynamics simulations themselves be carried out. Finally, the large amount of data produced by the simulations needs to be analyzed, interpreted, and set in relation to experimental evidence, if available.

16.4.1 Choice of the Electronic Structure Method

An electronic structure method is required in SH in order to find solutions of the electronic Schrödinger equation at every time step. These solutions then deliver energies, gradients, and (non-adiabatic, spin–orbit, or any other type of) couplings that are required to propagate the nuclear and electronic degrees of freedom in the SH algorithm.

The selection of the electronic structure method is actually one of the decisive choices that affects the outcome and feasibility of an SH project. This is because the electronic structure method affects the shape and accuracy of the PESs on which the dynamics is carried out, and inaccurate PESs will lead to systematically wrong dynamics. However, the electronic structure calculations also account for the vast majority of the computational cost, and therefore it is always necessary to find a compromise between accuracy and feasibility, especially when also aiming at a significant number of trajectories and/or a long simulation time.

The problem of choosing an electronic structure method is further complicated when electronic states of different characters need to be described simultaneously. This is because the different characters—valence states (e.g., $n \rightarrow \pi^*$, $\pi \rightarrow \pi^*$, $\pi \rightarrow \sigma^*$), Rydberg states, metal-centered ($d \rightarrow d$) states, charge transfer states, double excitations, etc.—often require different correlation treatments. Hence, finding a single method that describes all states in a balanced and reasonable way might not be trivial.

In many instances of non-adiabatic dynamics, the system evolves to situations where the ground state acquires a partial open-shell character, e.g., at near-degeneracies of ground state and an excited state, bond cleavage, or large molecular deformations. Often, these open-shell situations

cannot be described qualitatively correctly by a single Slater determinant, so that it is necessary to employ a multi-configurational method [82]. This is the reason why a significant amount of SH studies in the last decades have been based on the complete active space self-consistent field (CASSCF) method [83] (see Chapter 6). Its advantages are its multi-reference description, computational efficiency, and availability of analytical gradients and NAC vectors, but unfortunately in CASSCF statical and dynamical electron correlation are not well balanced. This can lead, e.g., to an overestimation of energies of $\pi\pi^*$ states by 1–3 eV. Such accuracy problems in CASSCF can be solved by employing a correlated multi-reference method, for example using multi-reference configuration interaction (MRCI, see Chapter 9) [84–86] or CAS perturbation theory (CASPT2, see Chapter 10) [87–89]. Both are among the most accurate excited-state electronic structure methods used nowadays, but their steep computational scaling restricts them to rather small molecules. Furthermore, for larger systems the accuracy of MRCI can be strongly deteriorated due to its lack of size-extensivity [82], whereas the results of CASPT2 depend on a number of empirical shift parameters (e.g., real/imaginary level shift [90], IPEA shift [91, 92]). Finally, for all multi-configurational and multi-reference methods one should keep in mind that accuracy and computational cost strongly depend on the size and choice of the active space, although in the future more elaborate active space schemes, like restricted active space SCF (RASSCF) or generalized active space SCF (GASSCF), might help alleviate these problems (see Chapter 6). Another important issue of these methods is the complexity of optimizing the employed wave functions, which might lead to reliability problems like bad convergence, undesired orbital rotations, discontinuous PESs, or intruder states.

Due to the limitations of multi-reference methods, single-reference methods have become very attractive for use in SH simulations, as they can deliver satisfactory accuracy at significantly lower computational cost [93]. Unfortunately, for the above-mentioned open-shell ground state situations, most single-reference methods will fail—either by not converging the ground state wave function or by producing unphysical S_1/S_0 conical intersection topologies. This strongly contrast to intersections between two excited states, which are generally described correctly with single-reference methods [94]. For this reason, it is usually recommended to employ single-reference methods only for dynamics which does not involve decay to the ground state [10].

Among the usable excited-state single-reference methods, (linear response) time-dependent density functional theory (TD-DFT) [95, 96] with or without the Tamm-Damcoff approximation (TDA), and also including the closely related time-dependent Hartree–Fock (TD–HF) and configuration interaction with singles (CIS), are likely the most popular electronic structure methods for SH simulations (see Chapter 2). Its attractiveness originates from its computational efficiency combined with the generally good accuracy of TD–DFT (at least when employing a suitable density functional). For a detailed discussion of the strengths and limitations of TD–DFT combined with SH, several recent publications can be recommended [97, 98] (see also Chapter 2).

A more recent alternative is correlated wave-function-based single-reference methods, most prominently the algebraic diagrammatic construction method, like the efficient ADC(2) or the more sophisticated ADC(3) [99, 100] (see Chapter 5). These methods have been shown [101] to yield accurate PESs for excited states, at a cost that is higher than TD–DFT but typically much lower than correlated multi-reference methods, and have been successfully employed in SH simulations [24, 102]. Another method that is closely related to ADC(2) is approximate coupled cluster (CC2) [103]. At approximately twice the cost of ADC(2), CC2 can deliver slightly more accurate results, owing to the fact that the ground state in CC2 is more strongly correlated. However, the Jacobian matrix in CC2 theory is non-Hermitian, which can lead to severe convergence problems whenever two excited states become degenerate [101]–making CC2 of limited use in SH

simulations. One exception is dynamics involving only S_1 and S_0, as it has been shown recently that CC2 can produce correct S_1/S_0 conical intersection topologies [94]. Another related way to obtain correlated single-reference wave functions is given by the family of equation-of-motion coupled cluster (EOM-CC) methods [104, 105] (see Chapter 4).

Some of the mentioned electronic structure methods–CASSCF, CASPT2, MRCI, CC2, ADC(2), and TD-DFT/TDA/TD–HF/CIS–constitute a set of *ab initio* methods that have been used for SH simulations. In order to compare these methods, Table 16.1 lists some of their strengths and limitations. The first part of the table shows which electronic situations can be covered with the methods. In the second part, we show which electronic quantities are available for these methods as of 2020. Besides the capabilities presented in the table, one also usually needs to consider how easily usable a method is—i.e., how much the method suffers from convergence problems, root switching, intruder states, required user effort, parametrization, etc.

Another, much cheaper, way in addition to ab initio methods is offered by semi-empirical electronic structure methods [27, 106, 107]. For excited states, most of the semi-empirical methods combine a form of CI or MRCI with a specifically parametrized Hamiltonian, for example OM*x* [106], AM3/PM3 [27], or some form of DFT [108]. These methods are typically orders of magnitude faster than ab initio methods, and can achieve reasonable accuracy if carefully parametrized. However, the parametrization is often not transferable, thus requiring reparametrization for each system. An independent method that shares the semi-empirical character and high efficiency is time-dependent density functional tight binding [109, 110], which is related to TD–DFT.

There are also a number of emerging electronic structure methods, which might be applied to SH simulations in the near future. Among the multi-reference methods, two interesting new approaches are the density matrix renormalization group [111, 112] (DMRG, see Chapter 7) technique and the Quantum Monte Carlo [113–115] (QMC, see Chapter 8) technique. Both methods allow the optimization of large multi-reference wave functions, e.g., with complete active spaces

Table 16.1 Qualitative overview over the capabilities of different electronic structure methods with respect to SH simulations.

	CASSCF	CASPT2	MRCI	CC2	ADC(2)	TD-DFT /TDA /TD–HF /CIS	TD-DFTB	semi-emp. CIS	semi-emp. MRCI
Static correlation	✓	✓	✓						✓
Dynamic correlation		✓	✓	✓	✓	✓	✓	✓	✓
Higher excitation ranks	✓	✓	✓						✓
S_m/S_n crossings	✓	✓	✓		✓	✓	✓	✓	✓
S_0/S_n crossings	✓	✓	✓	✓					✓
Dissociation	✓	✓	✓						✓
E+∇E	✓	✓	✓	✓	✓	✓	✓	✓	✓
NAC vectors	✓	✓	✓			✓			✓
Wave function overlaps	✓	✓	✓	✓	✓	✓	✓	✓	✓
SOC	✓	✓	✓	✓	✓	✓	✓	✓	✓
(T)DMs	✓	✓	✓	(✓)	✓	(✓)	(✓)	(✓)	✓

of at least 30 orbitals [116], opening up multi-reference methods for much more demanding systems than before. Among the single-reference methods, one example is the combination of the *GW* (Green's function *G* and screened Coulomb interaction *W*) method with the Bethe–Salpeter equation approach (*GW*/BSE) [117, 118], which is a popular method in solid-state physics but only recently started to be applied to molecular systems. Here, *GW* is a perturbative correction to a Kohn–Sham ground state [118], whereas the BSE is a Green's function ansatz to calculate excited states based on the *GW* ground state. As such, *GW*/BSE is a single-reference method with similar scaling to TD-DFT but a somewhat improved accuracy [118]. Other examples for emerging techniques focus on combining the advantages of single- and multi-reference methods (see Chapter 3). A general approach—applicable to different electronic structure methods like TD-DFT [119] or ADC [120]—is the spin–flip technique [121], where an open-shell triplet is used as the reference state and all singlet states—including S_0 are described consistently as spin–flip excitation from this reference. In this way, conical intersections between S_1 and S_0 can be described correctly [120]. Another technique is (state-average) spin-restricted ensemble-referenced Kohn–Sham (SA-)REKS [122, 123], which, loosely speaking, can be thought of as a DFT analogue of MCSCF theory.

A very important step in the preparation of SH simulations is a thorough validation of the electronic structure method. The first step is to check whether the vertical excitation energies at the S_0 minimum, calculated with the method of choice, are in agreement with the experimental absorption spectrum or with higher-level computations. However, a good agreement at this geometry does not necessarily mean that the electronic structure method will also describe energies properly outside of the Franck–Condon region. Hence, it is recommended that critical geometries are optimized—such as excited-state minima, minimum-energy crossing points, or transition states—that are relevant for the processes of interest. Then, the energies at these critical points and along paths connecting the critical points can be compared with a higher-level reference method to gain a more reliable estimation of the accuracy of the chosen method. For even more scrutiny, all relevant critical geometries should be found (independently) with the higher-level method, in order to ensure that the chosen method does not miss any important reaction path. To this end, a (possibly expensive but exhaustive) strategy is to run dynamics simulations with some initially chosen electronic structure method in order to explore the relevant configuration space. The results are then validated with higher-level single point calculations, and the electronic structure method for dynamics is appropriately adapted. Then, new dynamics simulations can be performed in a feedback loop until one is certain that the level of theory for dynamics is adequate.

16.4.2 Initial Conditions

In Section 16.3, we described in detail how in SH all nuclear and electronic properties can be propagated from a time step t to the next step, $t + \Delta t$. Of course, in order to set up such a trajectory, one has to provide the SH program with the values of all nuclear and electronic properties at $t = 0$. The set of values is called *initial conditions*.

According to Figure 16.2, the main nuclear and electronic variables that are propagated are: (i) nuclear positions $\vec{R}$, (ii) nuclear velocities $\vec{p}$, (iii) electronic wave function coefficients $\vec{c}$, and (iv) active electronic state β. The other quantities shown in Figure 16.2–acceleration/gradient, Hamiltonian matrix, and coupling matrix–directly depend on $\vec{R}$ and thus require no separate initial value. Hence, in order to compute all quantities at $t = 0$, one needs to define $\vec{R}$, $\vec{p}$, $\vec{c}$, as well as β, and then perform an electronic structure calculation at $\vec{R}$.

The initial values of $\vec{R}$, $\vec{p}$, $\vec{c}$, and β—the *initial conditions* in the narrower sense—are dependent on the chemical problem that one is trying to describe. For the most common case in SH simulations—excitation of a molecule in a thermally equilibrated ground state—$\vec{R}$ and $\vec{p}$ depend on the phase space distribution in the ground state, and $\vec{c}$, and β depend on the excitation process. There are two common techniques to randomly sample $\vec{R}$ and $\vec{p}$, which are shown in Figure 16.8(a) and b. In Wigner sampling one computes a (approximate) phase space probability distribution function representing the density of the ground state vibrational state, and then stochastically draws samples from this distribution. This has the advantage that quantum effects like zero-point energy can be adequately considered, but usually it is not trivial to find an appropriate distribution function. For polyatomic molecules, one usually resorts to a Wigner distribution of the solutions of a quantum harmonic oscillator in linear normal modes [124], which works well for small, stiff molecules.

An alternative to Wigner sampling is to run a long classical molecular dynamics simulation in the ground state, and randomly pick $(\vec{R}, \vec{p})$ pairs from this trajectory. This approach will not represent quantum effects like zero-point energy well, but does work well for large, poly-molecular systems with anharmonic or nonlinear modes and with multiple local minima in the ground state PES.

Once $(\vec{R}, \vec{p})$ pairs have been obtained by sampling, one can find the initial $\vec{c}$ and β. In the simplest case, one can use $c_\alpha(0) = \delta_{\alpha\beta}$, with β being one of the excited states. This β can be simply defined by the user (e.g., all trajectories start in S_2). A more appropriate and popular approach is to perform a single point calculation for each sampled $\vec{R}$ and compute a selection probability based on the obtained excitation energies and oscillator strengths of each state [125]:

$$p_\alpha(\vec{R}) = \frac{1}{p_{\text{norm}}} \frac{f_\alpha(\vec{R})}{|E_\alpha(\vec{R}) - E_{\text{init}}(\vec{R})|^2}, \tag{16.25}$$

where $E_{\text{init}}(\vec{R})$ is the energy of the initial state, f_α is the oscillator strength of state α, and p_{norm} is chosen such that the largest of all probabilities for all $\vec{R}$ is equal to unity. With these probabilities, one can stochastically select initial states. Additionally, often some restrictions on $|E_\alpha(\vec{R}) - E_{\text{init}}(\vec{R})|$ are imposed, in order to consider only a small excitation energy window. A more realistic alternative would be to explicitly simulate the interaction of a laser pulse with the molecule for every $\vec{R}$, although naturally this approach would be very expensive.

When preparing initial conditions, a very important thing to keep in mind is the size of the simulation ensemble, as SH simulations are basically stochastic samples of the photo-physical process. This entails at least three aspects: (i) the sampling of the initial phase space volume, (ii) the sampling of the initial active state and electronic coefficients, and (iii) the sampling of all hopping

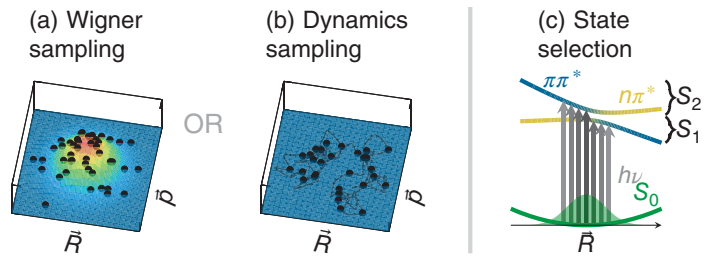


Figure 16.8 Overview of important steps in the initial condition generation. In (a) and (b), the two alternatives for phase space $(\vec{R}, \vec{p})$ sampling are shown. One of these techniques is required to generate the initial nuclear positions and velocities. In (c), the subsequent step of initial state selection is presented. This step is required to define the initial electronic coefficients and active state of the SH trajectories.

events during the trajectories. Hence, optimally one would prepare a large set of $(\vec{R}, \vec{p})$ pairs, for each pair one would consider all relevant initial β and $\vec{c}$, and for each $(\vec{R}, \vec{p}, \beta, \vec{c})$ one would prepare multiple trajectories that differ only in their random number sequence for hopping. In practice, one is usually limited in the number of trajectories one can compute by the availability of computational resources and the expense of the chosen electronic structure method, and therefore these sampling requirements can only be partially fulfilled.

16.4.3 Example Application and Trajectory Analysis

In order to explain possible ways of analyzing SH trajectories, in this section, we discuss a real-world example of a complete SH study. We will be investigating the excited-state dynamics of the methylene immonium cation $CH_2NH_2^+$ (also called the aminomethyl cation) in gas phase. This ion is isoelectronic to ethylene, but its positive charge reduces the influence of the Rydberg states that make simulations on ethylene very challenging [126]. The photo-induced dynamics of $CH_2NH_2^+$ was first simulated with SH in 2006 [127] at the CASSCF level of theory. Here, we will provide a description of the work flow for a SH study of this molecule with the MR-CISD(6,4)/aug-cc-pVDZ method for three singlet states.

On the left side of Figure 16.9, we sketch the work flow of preparation steps for the SH simulations. First, an optimization and frequency calculation is performed at the MR-CISD level of theory to obtain the ground state equilibrium geometry $\vec{R}_{eq}$, the vibrational frequencies $\{\omega_i\}$, and the normal mode vectors $\{\vec{n}_i\}$. With this information, it is possible to solve the quantum harmonic oscillator for the ground state. From the resulting Wigner distribution, we obtain the set of initial $(\vec{R}, \vec{p})$ pairs, where we generated 1000 initial conditions for this example. Subsequently, a vertical excitation calculation is carried out at every $\vec{R}_k$, yielding lists of excitation energies and oscillator strengths. Using this data and Gaussian convolution, one can simulate the absorption spectrum of the molecule, which is shown on the right of Figure 16.9.

For the initial state selection, we choose an excitation window with a width of 0.3 eV centered at the band maximum, specifically from 9.29 to 9.59 eV. In this way, all trajectories will start in S_2, as for all 1000 geometries the S_1 energy is outside this window. For the S_2, 233 geometries give

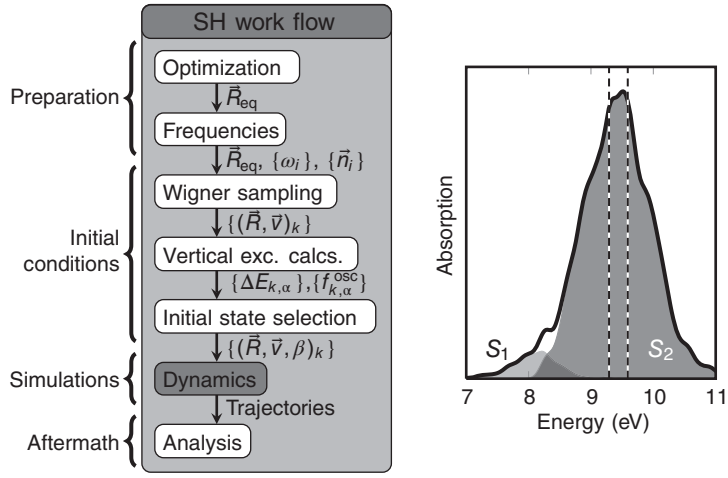


Figure 16.9 Scheme for a usual setup work flow of a SH simulation (left) and the simulated absorption spectrum of $CH_2NH_2^+$. The dashed lines indicate the energy window from where initial states were selected.

energies inside the window, and the stochastic selection process of Eq. (16.25) accepts 210 initial conditions.

In the next step, these 210 trajectories are propagated for 100 fs, using the MR-CISD method described above. Inspection of all trajectories will reveal that many of them display some violation of total energy conservation. Here, this is because the CAS(6,4) active space (π, π^*, σ_{CN}, σ^*_{CN}) is actually too small to describe all possible processes, especially the dissociation of H atoms. In these cases, large orbital rotations lead to discontinuities in the PESs and these lead to the total energy violations. Hence, for the remainder of this example, we have removed all problematic trajectories, so that for the statistical analysis 90 trajectories remain. In general, such a large number of problematic trajectories should instead be combated with an adjustment of the electronic structure method, e.g., enlargement of the active space.

In Figure 16.10, we present an overview of many results obtained from the analysis of the trajectories. In panel (a), we present an energy plot for a representative trajectory, showing the evolution from the initially occupied S_2 state at 9.6 eV. At 15 fs, the trajectory hops to the S_1 state, and after additional 45 fs to the ground state. The color of the plotted line indicates the oscillator strengths of the states, which is useful because in many cases the diabatic character can be learned from the oscillator strength, e.g., $\pi\pi^*$ states are bright and $n\pi^*$ states are dark. As discussed above, a single SH trajectory is of limited value, and instead statistical analysis should be carried out. In that sense, inspecting plots like the one in panel (a) is mostly useful to check all trajectories for computational artifacts and to formulate general hypotheses about the dynamics that can then be verified with statistics.

In panel (b) we show the evolution of the electronic populations in the ensemble of 90 trajectories. As can be seen, in $CH_2NH_2^+$ the initially excited S_2 state decays very quickly to the S_1, which subsequently decays to S_0. The S_0 population after 100 fs is about 80%. This is consistent with the total number of surface hops, shown in panel (c). As there are very few hops to the upper state, one can qualitatively describe the dynamics by a simple sequential reaction model, i.e., $S_2 \rightarrow S_1 \rightarrow S_0$. When the evolution of the populations is fit to the integrated rate laws of the kinetic model

$$N_{S_2}(t) = N_{S_2}(0) \exp\left(-\frac{t}{\tau_1}\right), \tag{16.26}$$

$$N_{S_1}(t) = N_{S_2}(0) \frac{\frac{1}{\tau_1}}{\frac{1}{\tau_1} - \frac{1}{\tau_2}} \left[\exp\left(-\frac{t}{\tau_2}\right) - \exp\left(-\frac{t}{\tau_1}\right)\right], \tag{16.27}$$

$$N_{S_0}(t) = N_{S_2}(0) \left(\frac{\frac{1}{\tau_2}}{\frac{1}{\tau_1} - \frac{1}{\tau_2}} \exp\left(-\frac{t}{\tau_1}\right) - \frac{\frac{1}{\tau_1}}{\frac{1}{\tau_1} - \frac{1}{\tau_2}} \exp\left(-\frac{t}{\tau_2}\right) + 1 \right), \tag{16.28}$$

it is possible to extract the time constants τ_1 and τ_2 from the fit. These give the decay time constant for the excited states, which in the example are 18 fs for S_2 and 51 fs for S_1. These time constants can then be compared to experimental results. For this purpose, it is also convenient to compute the sampling error [51, 58] of the time constants, which indicate whether the ensemble is large enough and the trajectories long enough to provide meaningful time constants. Together, panels (b) and (c) provide the most important information about the temporal evolution of the electronic wave function. Note, however, that in SH simulations one typically only has access to adiabatic populations, as the ones shown in (b), although diabatic populations are usually more useful because they directly relate to the electronic character and to experimental results. Obtaining diabatic populations from SH trajectories is possible in special cases [128, 129], but for polyatomic molecules in

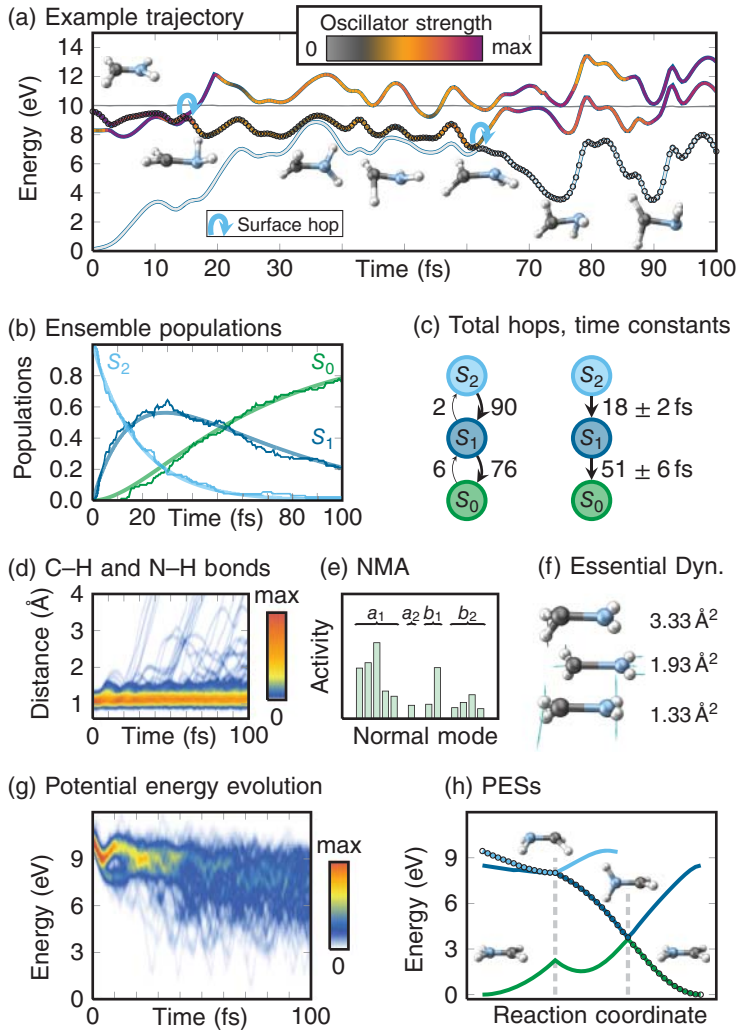


Figure 16.10 Collection of results from an ensemble of 90 SH trajectories for the $CH_2NH_2^+$ cation. The figure is discussed in the text.

general is very difficult. In this respect, automatic wave function analysis tools like the TheoDORE package [130] can be used to track the qualitative wave function character along a trajectory [28].

Besides the electronic evolution, SH trajectories also deliver a large amount of data about the nuclear degrees of freedom. In panel (d), we plot the evolution of the C–H and N–H bond length distribution. As can be seen, for these bonds no strong oscillations occur after excitation, showing that the bonds have similar lengths in the ground and excited state ($\pi\pi^*$). However, in some trajectories dissociation of a H atom seems to occur. Most probably, in these trajectories the electronic wave function character changed from $\pi\pi^*$ to a state involving a σ^* orbital, which leads to a repulsive PES. Of course, similar plots could be created for other degrees of freedom, depending on which process is most interesting to the researcher. Panels (e) and (f) show two techniques that help in finding important degrees of freedom in the dynamics. In normal mode analysis [131, 132], the nuclear motion is transformed to the ground state normal modes and, subsequently, the activity

in each normal mode is computed through statistical means. Normal modes with a higher activity, like in panel (e) the first three a_1 modes (scissoring and C=N stretch) or the second b_1 mode (pyramidalization at the carbon), can be regarded as more important for the dynamics. In essential dynamics analysis [133] (panel (f)), the Cartesian displacements from an average structure are statistically analyzed to find, directly, the most active degrees of freedom, i.e., not in terms of normal modes, but in terms of describing the dynamics with as few modes as possible via principal component analysis. In the example, this analysis yields as the strongest displacements the carbon pyramidalization (contributing a variance of 3.33 $Å^2$), followed by C=N stretch (1.93 $Å^2$) and the torsional mode (1.33 $Å^2$). Together, these three modes already account for 63% of the variance in the nuclear motion—showing that essential dynamics analysis can be used for effective dimensional reduction of the information in the trajectories.

Panel (g) shows how the potential energy of the trajectories evolves with time. It nicely presents how the ensemble reaches the S_2/S_1 crossing for the first time at 5 fs and splits into a branch staying on S_2 (>8.5 eV) and a branch that continues on S_1 (<8.5 eV). For later times, the upper branch again approaches the crossing around 20 fs, and eventually all trajectories assume a broad distribution of potential energy

Finally, panel (h) presents the relaxation mechanism found through the dynamics simulations—the hopping geometries served as starting points to optimize the conical intersections, and the plot in (h) was then obtained by linear interpolation in internal coordinates between the Franck–Condon point and these intersections. According to the scan, after excitation to S_2, the molecule should first pyramidalize to reach the S_2/S_1 crossing. Afterwards, torsion around the C=N double bond leads to the S_1/S_0 crossing and facilitates relaxation. Although this is a good qualitative description of the dynamics, we note that the simulated dynamics does not necessarily have to go through the optimized minimum-energy crossing points. In fact, panels (e) and (f) indicate that in the dynamics pyramidalization of the carbon atom is much stronger than pyramidalization of the nitrogen, whereas in panel (h) both atoms are pyramidalized at the S_2/S_1 intersection. This issue—that the real dynamics does not go exactly through the minimum of the intersection seam—should always be kept in mind when investigating the relaxation mechanism of a molecule.

16.5 Popular Implementations

Until the year 2020, the number of publicly available implementations of SH was relatively small, compared to the number of electronic structure codes. Hence, a brief mention of the most important SH codes is in order here, although it can be expected that this situation will have changed in the future. A more complete, up-to-date list of SH programs as of 2018 is given in Ref. [10]. In particular, that reference discusses that there are two basic types of implementations. In the first type, SH modules are provided as part of electronic structure packages, as e.g., in MOLCAS [134], TURBOMOLE [135], CPMD [97], or QChem [136]. This option is convenient for users, but is usually somewhat limited with regard to initial condition setup, ensemble management, and statistical analysis. The second type of implementation are the stand-alone SH packages with interfaces to electronic structure packages. Popular SH packages are ANT [137], JADE [41], Newton-X [40], PYXAID [138], and the SHARC package [11]. These packages are typically more complicated to operate, but usually offer large tool sets to perform automatized setup, management, and analysis tasks. They are also more likely to implement more sophisticated surface hopping algorithms like modern decoherence corrections, propagation schemes, or hopping methods.

16.6 Conclusion and Outlook

In this chapter we have reviewed the surface hopping method, which is a popular method for carrying out non-adiabatic dynamics simulations. Surface hopping is a mixed quantum-classical method, where electrons are described quantum mechanically and nuclei classically. The nuclear trajectories follow the potential energy surfaces obtained from solving the electronic Schrödinger equation, and non-adiabatic transitions between electronic states are included through instantaneous hops. Many independent trajectories are necessary to recover the branching into different possible relaxation pathways. The advantages of this approach is that due to the classical approximation for the nuclei large systems with hundreds of atoms can be treated, as long as the electronic Schrödinger equation can be approximately solved. Additionally, the independent trajectories make the surface hopping approach computationally efficient. Among the non-adiabatic dynamics methods presented in this volume, surface hopping is one of the cheapest methods. However, this efficiency comes at the price that surface hopping neglects a number of quantum effects, such as tunneling, nuclear interference, or zero-point energy.

The chapter comprehensively reviewed all parts of the surface hopping algorithm–nuclear motion, wave function propagation, decoherence correction, how to make the hopping decisions, nuclear momentum adjustment, and the inclusion of diverse coupling terms. We hope that these descriptions enable interested readers to choose appropriate algorithms for all these components of surface hopping, either when investigating the non-adiabatic dynamics of a chemical system, or when implementing their own surface hopping code.

A point that we did not highlight in this chapter, but which will surely grow continually in importance is the simulation of experimental observables based on surface hopping ensembles. Using appropriate methods, it is possible to simulate many different kinds of time-dependent spectra. For example, transient absorption spectra [139] can be computed from the transition dipole moments between active state and the other electronic states [140], whereas time-dependent photo-electron spectra [141] are accessible through photo-ionization cross sections [142–144]. Other spectroscopic methods that can, in principle, be computed, are time-resolved infrared spectroscopy [145], time-resolved X-Ray scattering [146] (see also Chapter 15), or time-resolved electron diffraction [147]. While these spectra simulations usually add to the computational cost of the surface hopping ensemble, they provide additional insight into the origin of the experimental signals and at the same time help validating the correctness of the surface hopping dynamics.

Acknowledgments

We thank Julia Westermayr and Anna Bäck for the data presented in Figure 16.10. Financial support from the University of Vienna and the Austrian Science Fund (project I2883) and computational support from the Vienna Scientific Cluster (VSC3) are gratefully acknowledged.

References

1 Tully, J.C. (1998). Mixed quantum-classical dynamics. *Faraday Discuss.* 110: 407–419.
2 Newton, I. (1726). *Philosophiae naturalis principia mathematica*. Cambridge: Harvard Univ. Press.

3 Doltsinis, N.L. (2006). Molecular dynamics beyond the Born–Oppenheimer approximation: mixed quantum-classical approaches. In: *Computational Nanoscience: Do It Yourself!*, volume 31 of *NIC Series* (eds. J. Grotendorst, S. Blügel and D. Marx), 389–409. Jülich: John von Neuman Institut for Computing.

4 Fabiano, E., Keal, T., and Thiel, W. (2008). Implementation of surface hopping molecular dynamics using semiempirical methods. *Chem. Phys.* 349: 334–347.

5 Barbatti, M. (2011). Nonadiabatic dynamics with trajectory surface hopping method. *WIREs Comput. Mol. Sci.* 1: 620–633.

6 Persico, M. and Granucci, G. (2014). An overview of non-adiabatic dynamics simulations methods, with focus on the direct approach versus the fitting of potential energy surfaces. *Theor. Chem. Acc.* 133: 1526.

7 Malhado, J.P., Bearpark, M.J., and Hynes, J.T. (2014). Non-adiabatic dynamics close to conical intersections and the surface hopping perspective. *Front. Chem.* 2: 97.

8 Wang, L., Akimov, A., and Prezhdo, O.V. (2016). Recent progress in surface hopping: 2011–2015. *J. Phys. Chem. Lett.* 7: 2100–2112.

9 Subotnik, J.E., Jain, A., Landry, B. et al. (2016). Understanding the surface hopping view of electronic transitions and decoherence. *Annu. Rev. Phys. Chem.* 67: 387–417.

10 Crespo-Otero, R. and Barbatti, M. (2018). Recent advances and perspectives on non-adiabatic mixed quantum–classical dynamics. *Chem. Rev.* 118: 7026–7068.

11 Mai, S., Marquetand, P., and González, L. (2018). Nonadiabatic dynamics: the SHARC approach. *WIREs Comput. Mol. Sci.* 8: e1370.

12 Tully, J.C. and Preston, R.K. (1971). Trajectory surface hopping approach to non-adiabatic molecular collisions: the reaction of H^+ with D_2. *J. Chem. Phys.* 55: 562–572.

13 Subotnik, J.E., Ouyang, W., and Landry, B.R. (2013). Can we derive Tully's surface-hopping algorithm from the semiclassical quantum Liouville equation? Almost, but only with decoherence. *J. Chem. Phys.* 139: 214107.

14 Kapral, R. (2016). Surface hopping from the perspective of quantum-classical Liouville dynamics. *Chem. Phys.* 481: 77–83.

15 Abedi, A., Agostini, F., Suzuki, Y., and Gross, E.K.U. (2013). Dynamical steps that bridge piecewise adiabatic shapes in the exact time-dependent potential energy surface. *Phys. Rev. Lett.* 110: 263001.

16 Agostini, F., Min, S.K., Abedi, A., and Gross, E.K.U. (2016). Quantum-classical non-adiabatic dynamics: coupled- vs independent-trajectory methods. *J. Chem. Theory Comput.* 12: 2127–2143.

17 Chen, H.T. and Reichman, D.R. (2016). On the accuracy of surface hopping dynamics in condensed phase non-adiabatic problems. *J. Chem. Phys.* 144: 094104.

18 Xie, W. and Domcke, W. (2017). Accuracy of trajectory surface-hopping methods: test for a two-dimensional model of the photodissociation of phenol. *J. Chem. Phys.* 147: 184114.

19 Gao, X. and Thiel, W. (2017). Non-Hermitian surface hopping. *Phys. Rev. E* 95: 013308.

20 Verlet, L. (1967). Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard–Jones molecules. *Phys. Rev.* 159: 98–103.

21 Hammes-Schiffer, S. and Tully, J.C. (1994). Proton transfer in solution: molecular dynamics with quantum transitions. *J. Chem. Phys.* 101: 4657–4667.

22 Pittner, J., Lischka, H., and Barbatti, M. (2009). Optimization of mixed quantum-classical dynamics: time-derivative coupling terms and selected couplings. *Chem. Phys.* 356: 147–152.

23 Tapavicza, E., Tavernelli, I., and Rothlisberger, U. (2007). Trajectory surface hopping within linear response time-dependent density-functional theory. *Phys. Rev. Lett.* 98: 023001.

24 Plasser, F., Crespo-Otero, R., Pederzoli, M. et al. (2014). Surface hopping dynamics with correlated single-reference methods: 9H-adenine as a case study. *J. Chem. Theory Comput.* 10: 1395–1405.

25 Plasser, F., Ruckenbauer, M., Mai et al. (2016). Efficient and flexible computation of many-electron wave function overlaps. *J. Chem. Theory Comput.* 12: 1207.

26 Fernandez-Alberti, S., Roitberg, A.E., Nelson, T., and Tretiak, S. (2012). Identification of unavoided crossings in non-adiabatic photoexcited dynamics involving multiple electronic states in polyatomic conjugated molecules. *J. Chem. Phys.* 137: 014512.

27 Granucci, G., Persico, M., and Toniolo, A. (2001). Direct semiclassical simulation of photochemical processes with semiempirical wave functions. *J. Chem. Phys.* 114: 10608–10615.

28 Plasser, F., Granucci, G., Pittner, J. et al. (2012). Surface hopping dynamics using a locally diabatic formalism: charge transfer in the ethylene dimer cation and excited state dynamics in the 2-pyridone dimer. *J. Chem. Phys.* 137: 22A514.

29 Tully, J.C. (1990). Molecular dynamics with electronic transitions. *J. Chem. Phys.* 93: 1061–1071.

30 Schwartz, B.J., Bittner, E.R., Prezhdo, O.V., and Rossky, P.J. (1996). Quantum decoherence and the isotope effect in condensed phase non-adiabatic molecular dynamics simulations. *J. Chem. Phys.* 104: 5942–5955.

31 Prezhdo, O.V. and Rossky, P.J. (1997). Evaluation of quantum transition rates from quantum-classical molecular dynamics simulations. *J. Chem. Phys.* 107: 5863–5878.

32 Cheng, S.C., Zhu, C., Liang, K.K. et al. (2008). Algorithmic decoherence time for decay-of-mixing non-Born–Oppenheimer dynamics. *J. Chem. Phys.* 129: 024112.

33 Granucci, G., Persico, M., and Zoccante, A. (2010). Including quantum decoherence in surface hopping. *J. Chem. Phys.* 133: 134111.

34 Ben-Nun, M., Quenneville, J., and Martínez, T.J. (2000). Ab initio multiple spawning: photochemistry from first principles quantum molecular dynamics. *J. Phys. Chem. A* 104: 5161–5175.

35 Makhov, D.V., Glover, W.J., Martinez, T.J., and Shalashilin, D.V. (2014). Ab initio multiple cloning algorithm for quantum non-adiabatic molecular dynamics. *J. Chem. Phys.* 141: 054110.

36 Fang, J.Y. and Hammes-Schiffer, S. (1999). Improvement of the internal consistency in trajectory surface hopping. *J. Phys. Chem. A* 103: 9399–9407.

37 Schwartz, B.J. and Rossky, P.J. (1994). Aqueous solvation dynamics with a quantum mechanical solute: computer simulation studies of the photoexcited hydrated electron. *J. Chem. Phys.* 101: 6902–6916.

38 Zhu, C., Nangia, S., Jasper, A.W., and Truhlar, D.G. (2004). Coherent switching with decay of mixing: an improved treatment of electronic coherence for non-Born–Oppenheimer trajectories. *J. Chem. Phys.* 121: 7658–7670.

39 Granucci, G. and Persico, M. (2007). Critical appraisal of the fewest switches algorithm for surface hopping. *J. Chem. Phys.* 126: 134114.

40 Barbatti, M., Ruckenbauer, M., Plasser, F. et al. (2014). Newton-X: a surface-hopping program for non-adiabatic molecular dynamics. *WIREs Comput. Mol. Sci.* 4: 26–33.

41 Du, L. and Lan, Z. (2015). An on-the-fly surface-hopping program JADE for non-adiabatic molecular dynamics of polyatomic systems: implementation and applications. *J. Chem. Theory Comput.* 11: 1360–1374.

42 Subotnik, J.E. and Shenvi, N. (2011). A new approach to decoherence and momentum rescaling in the surface hopping algorithm. *J. Chem. Phys.* 134: 024105.

43 Jain, A., Alguire, E., and Subotnik, J.E. (2016). An efficient, augmented surface hopping algorithm that includes decoherence for use in large-scale simulations. *J. Chem. Theory Comput.* 12: 5256–5268.

44 Ha, J.K., Lee, I.S., and Min, S.K. (2018). Surface hopping dynamics beyond non-adiabatic couplings for quantum coherence. *J. Phys. Chem. Lett.* 9: 1097–1104.

45 Zhu, C. and Nakamura, H. (1992). The two-state linear curve crossing problems revisited. II. Analytical approximations for the Stokes constant and scattering matrix: the Landau–Zener case. *J. Chem. Phys.* 97: 8497–8514.

46 Nakamura, H. (2002). *Nonadiabatic Transition: Concepts, Basic Theories and Applications*. World Scientific.

47 Ishida, T., Nanbu, S., and Nakamura, H. (2017). Clarification of non-adiabatic chemical dynamics by the Zhu–Nakamura theory of non-adiabatic transition: from tri-atomic systems to reactions in solutions. *Int. Rev. Phys. Chem.* 36: 229–285.

48 Mai, S., Marquetand, P., and González, L. (2015). A general method to describe intersystem crossing dynamics in trajectory surface hopping. *Int. J. Quantum Chem.* 115: 1215–1231.

49 Lisinetskaya, P. and Mitrić, R. (2011). Simulation of laser-induced coupled electron-nuclear dynamics and time-resolved harmonic spectra in complex systems. *Phys. Rev. A* 83: 033408.

50 Wang, L., Trivedi, D., and Prezhdo, O.V. (2014). Global flux surface hopping approach for mixed quantum-classical dynamics. *J. Chem. Theory Comput.* 10: 3598–3605.

51 Nangia, S., Jasper, A.W., Miller, T.F., and Truhlar, D.G. (2004). Army ants algorithm for rare event sampling of delocalized non-adiabatic transitions by trajectory surface hopping and the estimation of sampling errors by the bootstrap method. *J. Chem. Phys.* 120: 3586–3597.

52 Jaeger, H.M., Fischer, S., and Prezhdo, O.V. (2012). Decoherence-induced surface hopping. *J. Chem. Phys.* 137: 22A545.

53 Wang, L., Sifain, A.E., and Prezhdo, O.V. (2015). Fewest switches surface hopping in Liouville space. *J. Phys. Chem. Lett.* 6: 3827–3833.

54 Barbatti, M., Lan, Z., Crespo-Otero, R. et al. (2012). Critical appraisal of excited state non-adiabatic dynamics simulations of 9H-adenine. *J. Chem. Phys.* 137: 22A503.

55 Nakayama, A., Arai, G., Yamazaki, S., and Taketsugu, T. (2013). Solvent effects on the ultrafast nonradiative deactivation mechanisms of thymine in aqueous solution: excited-state QM/MM molecular dynamics simulations. *J. Chem. Phys.* 139: 214304.

56 Stojanović, L., Bai, S., Nagesh, J. et al. (2016). New insights into the state trapping of UV-excited thymine. *Molecules* 21: 1603.

57 Müller, U. and Stock, G. (1997). Surface-hopping modeling of photoinduced relaxation dynamics on coupled potential-energy surfaces. *J. Chem. Phys.* 107: 6230–6245.

58 Mai, S., Marquetand, P., and González, L. (2016). Intersystem crossing pathways in the non-canonical nucleobase 2-thiouracil: a time-dependent picture. *J. Phys. Chem. Lett.* 7: 1978–1983.

59 Pechukas, P. (1969). Time-dependent semiclassical scattering theory. II. Atomic collisions. *Phys. Rev.* 181: 174–185.

60 Tully, J.C. (1991). Nonadiabatic molecular dynamics. *Int. J. Quantum Chem.* 40: 299–309.

61 Herman, M.F. (1984). Nonadiabatic semiclassical scattering. I. Analysis of generalized surface hopping procedures. *J. Chem. Phys.* 81: 754–763.

62 Coker, D.F. and Xiao, L. (1995). Methods for molecular dynamics with non-adiabatic transitions. *J. Chem. Phys.* 102: 496–510.

63 Hack, M.D., Jasper, A.W., Volobuev, Y.L. et al. (1999). Quantum mechanical and quasiclassical trajectory surface hopping studies of the electronically non-adiabatic predissociation of the $\tilde{A}$ state of NaH_2. *J. Phys. Chem. A* 103: 6309–6326.

64 Jasper, A.W., Stechmann, S.N., and Truhlar, D.G. (2002). Fewest-switches with time uncertainty: a modified trajectory surface-hopping algorithm with better accuracy for classically forbidden electronic transitions. *J. Chem. Phys.* 116: 5424–5431.

65 Jasper, A.W., Hack, M.D., and Truhlar, D.G. (2001). The treatment of classically forbidden electronic transitions in semiclassical trajectory surface hopping calculations. *J. Chem. Phys.* 115: 1804–1816.

66 Jasper, A.W. and Truhlar, D.G. (2003). Improved treatment of momentum at classically forbidden electronic transitions in trajectory surface hopping calculations. *Chem. Phys. Lett.* 369: 60–67.

67 Reiher, M. and Wolf, A. (2009). *Relativistic Quantum Chemistry*. Wiley VCH Verlag Weinheim.

68 Heß, B.A., Marian, C.M., Wahlgren, U., and Gropen, O. (1996). A mean-field spin–orbit method applicable to correlated wave functions. *Chem. Phys. Lett.* 251: 365–371.

69 Malmqvist, P.Å., Roos, B.O., and Schimmelpfennig, B. (2002). The restricted active space (RAS) state interaction approach with spin–orbit coupling. *Chem. Phys. Lett.* 357: 230.

70 Neese, F. (2005). Efficient and accurate approximations to the molecular spin–orbit coupling operator and their use in molecular *g*-tensor calculations. *J. Chem. Phys.* 122: 034107.

71 Neese, F. (2012). The ORCA program system. *WIREs Comput. Mol. Sci.* 2: 73–78.

72 Szabó, P. and Gustafsson, M. (2017). A surface-hopping method for semiclassical calculations of cross sections for radiative association with electronic transitions. *J. Chem. Phys.* 147: 094308.

73 Granucci, G., Persico, M., and Spighi, G. (2012). Surface hopping trajectory simulations with spin–orbit and dynamical couplings. *J. Chem. Phys.* 137: 22A501.

74 Herman, M.F. (1999). Choosing a good representation of the quantum state wave functions for semiclassical surface hopping calculations. *J. Chem. Phys.* 111: 10427–10435.

75 Fiedlschuster, T., Handt, J., Gross, E.K.U., and Schmidt, R. (2017). Surface hopping in laser-driven molecular dynamics. *Phys. Rev. A* 95: 063424.

76 Belcher, L.T. (2011). Gradients and non-adiabatic derivative coupling terms for spin–orbit wave functions, Ph.D. thesis, Air Force Institute of Technology, Department of the Air Force, Air University.

77 Granucci, G. and Persico, M. (2011). Gradients for configuration interaction energies with spin–orbit coupling in a semiempirical framework. *J. Comput. Chem.* 32: 2690–2696.

78 Pederzoli, M. and Pittner, J. (2017). A new approach to molecular dynamics with non-adiabatic and spin–orbit effects with applications to QM/MM simulations of thiophene and selenophene. *J. Chem. Phys.* 146: 114101.

79 Vallet, V., Maron, L., Teichteil, C., and Flament, J.P. (2000). A two-step uncontracted determinantal effective Hamiltonian-based SO-CI method. *J. Chem. Phys.* 113: 1391–1402.

80 Mai, S., Müller, T., Marquetand, P. et al. (2014). Perturbational treatment of spin–orbit coupling for generally applicable high-level multi-reference methods. *J. Chem. Phys.* 141: 074105.

81 Wang, F. and Ziegler, T. (2005). A simplified relativistic time-dependent density-functional theory formalism for the calculations of excitation energies including spin–orbit coupling effect. *J. Chem. Phys.* 123: 154102.

82 Szalay, P.G., Müller, T., Gidofalvi, G. et al. (2012). Multiconfiguration self-consistent field and multireference configuration interaction methods and applications. *Chem. Rev.* 112: 108–181.

83 Roos, B.O., Taylor, P.R., and Siegbahn, P.E.M. (1980). A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach. *Chem. Phys.* 48: 157–173.

84 Shepard, R., Lischka, H., Szalay, P.G. et al. (1992). A general multireference configuration interaction gradient program. *J. Chem. Phys.* 96: 2085–2098.

85 Lischka, H., Dallos, M., and Shepard, R. (2002). Analytic MRCI gradient for excited states: formalism and application to the n-π^* valence- and n-(3s,3p) Rydberg states of formaldehyde. *Mol. Phys.* 100: 1647–1658.

86 Lischka, H., Dallos, M., Szalay, P.G. et al. (2004). Analytic evaluation of non-adiabatic coupling terms at the MR-CI level. I. Formalism. *J. Chem. Phys.* 120: 7322–7329.

87 Andersson, K., Malmqvist, P.Å., Roos, B.O. et al. (1990). Second-order perturbation theory with a CASSCF reference function. *J. Phys. Chem.* 94: 5483–5488.

88 MacLeod, M.K. and Shiozaki, T. (2015). Communication: automatic code generation enables nuclear gradient computations for fully internally contracted multireference theory. *J. Chem. Phys.* 142: 051103.

89 Park, J.W. and Shiozaki, T. (2017). On-the-fly CASPT2 surface-hopping dynamics. *J. Chem. Theory Comput.* 13: 3676–3683.

90 Forsberg, N. and Malmqvist, P.Å. (1997). Multiconfiguration perturbation theory with imaginary level shift. *Chem. Phys. Lett.* 274: 196–204.

91 Ghigo, G., Roos, B.O., and Malmqvist, P.Å. (2004). A modified definition of the zeroth-order hamiltonian in multiconfigurational perturbation theory (CASPT2). *Chem. Phys. Lett.* 396: 142–149.

92 Zobel, J.P., Nogueira, J.J., and González, L. (2017). The IPEA dilemma in CASPT2. *Chem. Sci.* 8: 1482–1499.

93 Dreuw, A. and Head-Gordon, M. (2005). Single-reference ab initio methods for the calculation of excited states of large molecules. *Chem. Rev.* 105: 4009–4037.

94 Tuna, D., Lefrancois, D., Wolański, Ł. et al. (2015). Assessment of approximate coupled-cluster and algebraic-diagrammatic-construction methods for ground- and excited-state reaction paths and the conical-intersection seam of a retinal-chromophore model. *J. Chem. Theory Comput.* 11: 5758–5781.

95 Casida, M. and Huix-Rotllant, M. (2012). Progress in time-dependent density-functional theory. *Annu. Rev. Phys. Chem.* 63: 287–323.

96 Tavernelli, I., Tapavicza, E., and Rothlisberger, U. (2009). Nonadiabatic coupling vectors within linear response time-dependent density functional theory. *J. Chem. Phys.* 130: 124107.

97 Curchod, B.F.E., Rothlisberger, U., and Tavernelli, I. (2013). Trajectory-based non-adiabatic dynamics with time-dependent density functional theory. *ChemPhysChem* 14: 1314–1340.

98 Barbatti, M. and Crespo-Otero, R. (2016). *Density-Functional Methods for Excited States, Chapter "Surface Hopping Dynamics with DFT Excited States"*, 415–444. Cham: Springer International Publishing.

99 Trofimov, A.B. and Schirmer, J. (1995). An efficient polarization propagator approach to valence electron excitation spectra. *J. Phys. B Atomic Mol. Phys.* 28: 2299.

100 Dreuw, A. and Wormit, M. (2015). The algebraic diagrammatic construction scheme for the polarization propagator for the calculation of excited states. *WIREs Comput. Mol. Sci.* 5: 82–95.

101 Hättig, C. (2005). Structure optimizations for excited states with correlated second-order methods: CC2 and ADC(2). *Adv. Quantum Chem.* 50: 37–60.

102 Mai, S., Plasser, F., Pabst, M. et al. (2017). Surface hopping dynamics including intersystem crossing using the algebraic diagrammatic construction method. *J. Chem. Phys.* 147: 184109.

103 Christiansen, O., Koch, H., and Jørgensen, P. (1995). The second-order approximate coupled cluster singles and doubles model CC2. *Chem. Phys. Lett.* 243: 409–418.

104 Krylov, A.I. (2008). Equation-of-motion coupled-cluster methods for open-shell and electronically excited species: the Hitchhiker's guide to Fock space. *Annu. Rev. Phys. Chem.* 59: 433–462.

105 Bartlett, R.J. (2011). Coupled-cluster theory and its equation-of-motion extensions. *WIREs Comput. Mol. Sci.* 2: 126–138.

106 Thiel, W. (2014). Semiempirical quantum-chemical methods. *WIREs Comput. Mol. Sci.* 4: 145–157.

107 Nelson, T., Fernandez-Alberti, S., Roitberg, A.E., and Tretiak, S. (2014). Nonadiabatic excited-state molecular dynamics: modeling photophysics in organic conjugated materials. *Acc. Chem. Res.* 47: 1155–1164.

108 Grimme, S. and Waletzke, M. (1999). A combination of Kohn–Sham density functional theory and multi-reference configuration interaction methods. *J. Chem. Phys.* 111: 5645–5655.

109 Seifert, G. and Joswig, J.O. (2012). Density-functional tight binding—an approximate density-functional theory method. *WIREs Comput. Mol. Sci.* 2: 456–465.

110 Humeniuk, A. and Mitrić, R. (2017). DFTBaby: a software package for non-adiabatic molecular dynamics simulations based on long-range corrected tight-binding TD-DFT(B). *Comput. Phys. Commun.* 221: 174–202.

111 Marti, K.H. and Reiher, M. (2010). The density matrix renormalization group algorithm in quantum chemistry. *Z. Phys. Chem.* 224: 583.

112 Chan, G.K.L. and Sharma, S. (2011). The density matrix renormalization group in quantum chemistry. *Annu. Rev. Phys. Chem.* 62: 465–481.

113 Filippi, C. and Umrigar, C.J. (1996). Multiconfiguration wave functions for quantum Monte Carlo calculations of first-row diatomic molecules. *J. Chem. Phys.* 105: 213–226.

114 Austin, B.M., Zubarev, D.Y., and Lester, W.A. (2012). Quantum Monte Carlo and related approaches. *Chem. Rev.* 112: 263–288.

115 Booth, G.H., Grüneis, A., Kresse, G., and Alavi, A. (2013). Towards an exact description of electronic wave functions in real solids. *Nature* 493: 365–370.

116 Li Manni, G., Smart, S.D., and Alavi, A. (2016). Combining the complete active space self-consistent field method and the full configuration interaction quantum Monte Carlo within a super-CI framework, with application to challenging metal-porphyrins. *J. Chem. Theory Comput.* 12: 1245–1258.

117 Leng, X., Jin, F., Wei, M., and Ma, Y. (2016). GW method and Bethe–Salpeter equation for calculating electronic excitations. *WIREs Comput. Mol. Sci.* 6: 532–550.

118 Jacquemin, D., Duchemin, I., and Blase, X. (2017). Is the Bethe-Salpeter formalism accurate for excitation energies? Comparisons with TD-DFT, CASPT2, and EOM-CCSD. *J. Phys. Chem. Lett.* 8: 1524–1529.

119 Seth, M., Mazur, G., and Ziegler, T. (2011). Time-dependent density functional theory gradients in the Amsterdam density functional package: geometry optimizations of spin-flip excitations. *Theor. Chem. Acc.* 129: 331–342.

120 Lefrancois, D., Tuna, D., Martínez, T.J., and Dreuw, A. (2017). The spin-flip variant of the algebraic-diagrammatic construction yields the correct topology of S_1/S_0 conical intersections. *J. Chem. Theory Comput.* 13: 4436–4441.

121 Gagliardi, L., Truhlar, D.G., Li Manni, G. et al. (2017). Multiconfiguration pair-density functional theory: a new way to treat strongly correlated systems. *Acc. Chem. Res.* 50: 66–73.

122 Filatov, M. and Shaik, S. (1999). A spin-restricted ensemble-referenced Kohn–Sham method and its application to diradicaloid situations. *Chem. Phys. Lett.* 304: 429–437.

123 Huix-Rotllant, M., Filatov, M., Gozem, S. et al. (2013). Assessment of density functional theory for describing the correlation effects on the ground and excited state potential energy surfaces of a retinal chromophore model. *J. Chem. Theory Comput.* 9: 3917–3932.

124 Barbatti, M. and Sen, K. (2016). Effects of different initial condition samplings on photodynamics and spectrum of pyrrole. *Int. J. Quantum Chem.* 116: 762–771.

125 Barbatti, M., Granucci, G., Persico, M. et al. (2007). The on-the-fly surface-hopping program system Newton-X: application to ab initio simulation of the non-adiabatic photodynamics of benchmark systems. *J. Photochem. Photobiol. A* 190: 228–240.

126 Sellner, B., Barbatti, M., Müller, T. et al. (2013). Ultrafast non-adiabatic dynamics of ethylene including Rydberg states. *Mol. Phys.* 111: 2439–2450.

127 Barbatti, M., Aquino, A.J.A., and Lischka, H. (2006). Ultrafast two-step process in the non-adiabatic relaxation of the $CH_2NH_2^+$ molecule. *Mol. Phys.* 104: 1053–1060.

128 Landry, B.R., Falk, M.J., and Subotnik, J.E. (2013). Communication: the correct interpretation of surface hopping trajectories: how to calculate electronic properties. *J. Chem. Phys.* 139: 211101.

129 Mai, S., Marquetand, P., and González, L. (2014). Non-adiabatic dynamics in SO_2: II. The role of triplet states studied by surface-hopping simulations. *J. Chem. Phys.* 140: 204302.

130 Plasser, F. (2017). TheoDORE: a package for theoretical density, orbital relaxation, and exciton analysis, http://theodore-qc.sourceforge.net.

131 Kurtz, L., Hofmann, A., and de Vivie-Riedle, R. (2001). Ground state normal mode analysis: linking excited state dynamics and experimental observables. *J. Chem. Phys.* 114: 6151–6159.

132 Plasser, F., Barbatti, M., Aquino, A.J.A., and Lischka, H. (2009). Excited-state diproton transfer in [2,2'-bipyridyl]-3,3'-diol: the mechanism is sequential, not concerted. *J. Phys. Chem. A* 113: 8490–8499.

133 Amadei, A., Linssen, A.B.M., and Berendsen, H.J.C. (1993). Essential dynamics of proteins. *Proteins* 17: 412–425.

134 Aquilante, F., Autschbach, J., Carlson, R.K. et al. (2015). Molcas 8: new capabilities for multiconfigurational quantum chemical calculations across the periodic table. *J. Comput. Chem.* 37: 506–541.

135 Furche, F., Ahlrichs, R., Hättig, C. et al. (2014). Turbomole. *WIREs Comput. Mol. Sci.* 4: 91–100.

136 Krylov, A.I. and Gill, P.M. (2013). Q-Chem: an engine for innovation. *WIREs Comput. Mol. Sci.* 3: 317–326.

137 Zheng, J., Li, Z.H., Jasper, A.W. et al. (2017). Ant, version 2017, University of Minnesota, Minneapolis, http://comp.chem.umn.edu/ant.

138 Akimov, A.V. and Prezhdo, O.V. (2013). The PYXAID program for non-adiabatic molecular dynamics in condensed matter systems. *J. Chem. Theory Comput.* 9: 4959–4972.

139 Berera, R., Grondelle, R., and Kennis, J.T.M. (2009). Ultrafast transient absorption spectroscopy: principles and application to photosynthetic systems. *Photosynth. Res.* 101: 105–118.

140 Petit, A.S. and Subotnik, J.E. (2014). Calculating time-resolved differential absorbance spectra for ultrafast pump-probe experiments with surface hopping trajectories. *J. Chem. Phys.* 141: 154108.

141 Stolow, A. and Underwood, J.G. (2008). Time-resolved photoelectron spectroscopy of non-adiabatic dynamics in polyatomic molecules. In: *Advances in Chemical Physics* (ed. S.A. Rice), 497–584. John Wiley & Sons, Inc.

142 Hudock, H.R., Levine, B.G., Thompson, A.L. et al. (2007). Ab initio molecular dynamics and time-resolved photoelectron spectroscopy of electronically excited uracil and thymine. *J. Phys. Chem. A* 111: 8500–8508.

143 Mitrić, R., Petersen, J., Wohlgemuth, M. et al. (2011). Time-resolved femtosecond photoelectron spectroscopy by field-induced surface hopping. *J. Phys. Chem. A* 115: 3755–3765.

144 Arbelo-González, W., Crespo-Otero, R., and Barbatti, M. (2016). Steady and time-resolved photoelectron spectra based on nuclear ensembles. *J. Chem. Theory Comput.* 12: 5037–5049.

145 Nibbering, E.T., Fidder, H., and Pines, E. (2005). Ultrafast chemistry: using time-resolved vibrational spectroscopy for interrogation of structural dynamics. *Annu. Rev. Phys. Chem.* 56: 337–367.

146 Bressler, C. and Chergui, M. (2010). Molecular structural dynamics probed by ultrafast X-ray absorption spectroscopy. *Annu. Rev. Phys. Chem.* 61: 263–282.

147 Sciaini, G. and Miller, R.J.D. (2011). Femtosecond electron diffraction: Heralding the era of atomically resolved dynamics. *Rep. Prog. Phys.* 74: 096101.

17

Exact Factorization of the Electron–Nuclear Wave Function: Theory and Applications

Federica Agostini[1] *and E. K. U. Gross*[2]

[1] *Université Paris-Saclay, CNRS, Institut de Chimie Physique UMR8000, 91405, Orsay, France*
[2] *Fritz Haber Center for Molecular Dynamics, Institute of Chemistry, The Hebrew University of Jerusalem, Jerusalem 91904, Israel*

Abstract

In this Chapter we review the exact factorization of the electron-nuclear wave function. The molecular wave function, solution of a time-dependent Schröodinger equation, is factored into a nuclear wave function and an electronic wave function with parametric dependence on nuclear configuration. This factorization resembles the (approximate) adiabatic product of a single Born-Oppenheimer state and a time-dependent nuclear wave packet, but it introduces a fundamental difference: both terms of the product are explicitly time-dependent.

Such feature introduces new concepts of time-dependent vector potential and time-dependent potential energy surface that allow for the treatment of nonadiabatic dynamics, thus of dynamics beyond the Born-Oppenheimer approximation. The theoretical framework of the exact factorization is presented, also in connection to the more standard Born-Huang (still exact) representation of the molecular wave function. A trajectory-based approach to nonadiabatic dynamics is derived from the exact factorization. A discussion on the connection between the molecular Berry phase and the corresponding quantity arising from the exact factorization is briefly discussed.

17.1 Introduction

Ab initio molecular dynamics is nowadays a cornerstone of the fields of Theoretical Chemistry and Chemical Physics. Numerical simulations allow us to achieve a profound understanding of structural and dynamical properties of matter at the microscopic level, for a large variety of systems, from isolated molecules, or liquids, to proteins, and a large variety of phenomena, from equilibrium processes to ultrafast photo-activated relaxation. Ab initio molecular dynamics relies on some approximate solution of the time-dependent Schrödinger equation for systems of electrons and nuclei (see for instance Refs. [1–7]). Approximations are indeed necessary since the exponential scaling of the (exact) quantum-mechanical problem with the number of degrees of freedom makes its numerical treatment prohibitive, for all but the smallest systems.

In 1927, Born and Oppenheimer [8] posed the basis for what is currently the standard way of visualizing motion in molecular systems: the electrons generate the potential(s) felt by the nuclei during their evolution. This picture is retained either within the adiabatic, Born–Oppenheimer,

Quantum Chemistry and Dynamics of Excited States: Methods and Applications,
First Edition. Edited by Leticia González and Roland Lindh.

approximation, where the electrons are "statically" found in a single eigenstate, or in non-adiabatic situations, when multiple excited states are populated and can exchange amplitude during the dynamics. Although this perspective is perhaps the most adequate for practical applications in combination with quantum chemistry, alternative viewpoints, such as the pre-Born–Oppenheimer molecular structure theory [9–11], have been proposed and are presently being developed.

The Born–Oppenheimer perspective of coupled electronic and nuclear motion in molecules relies on the concept of electronic (adiabatic) potential energy surfaces. It appears to be very useful for developing approximations to the quantum-dynamical problem: those potentials, in fact, can be used to guide the classical-like evolution of the nuclei. Perhaps the most successful approximation used in the field of quantum molecular dynamics simulations is the use of – quantum, semiclassical, or classical – trajectories to sample the nuclear configuration space and to mimic the quantum-mechanical nuclear motion. Within the adiabatic approximation, when a single electronic eigenstate (usually the ground state), and the corresponding potential energy surface, is populated, trajectories are generated according to the force determined from the single potential that appears in the problem. Clearly, the situation becomes more involved in non-adiabatic circumstances, that is when excited states are also available during the dynamics. In those cases, a generally applied procedure to guide nuclear trajectories does not yet exist, due to the fact that it is not possible to identify in this framework *the* potential determining the force. Solutions to this problem can be envisaged, and currently include surface hopping methods [4, 12] (where trajectories can hop between adiabatic surfaces), Ehrenfest dynamics [3, 13] (where trajectories follow a mean-field potential), or multiple spawning approaches [2, 14] (where trajectory-basis functions are spawned in regions of strong coupling between states and are coupled together). However, it seems that the Born–Oppenheimer framework cannot help to easily, and uniquely, answer the question: What force governs the classical-like evolution of the nuclei in non-adiabatic dynamics?

In addition to these issues arising in the Born–Oppenheimer framework, conical intersections should also be mentioned as curious, and problematic, features of the adiabatic potential energy surfaces. Conical intersections are often invoked to interpret relaxation processes undergone by photoexcited molecules. They are typical examples of the breakdown of the Born–Oppenheimer approximation, and represent efficient funnels [15–17] for population transfer between electronic states, mediated by nuclear motion. They are regions of configuration space where the adiabatic potential energy surfaces are degenerate – thus producing a diverging coupling between the corresponding electronic states – and exhibit, within the so-called branching space, a double-cone shape. In the literature, conical intersections have been studied not only for their role in non-adiabatic processes, but also for the effect that the related Berry phase has on adiabatic phenomena [17–33]. The concept of conical intersections is based on the Born–Oppenheimer representation of the electron–nuclear problem. Despite the fundamental and numerical challenges encountered when dealing with conical intersections, such representation is still commonly employed to perform non-adiabatic ab initio molecular dynamics simulations.

In this chapter, an alternative perspective on the electron–nuclear time-dependent problem will be proposed and analyzed: the exact factorization of the electron–nuclear wave function [34–37]. In the exact-factorization framework, the issues related to the Born–Oppenheimer perspective raised in the previous paragraphs can be easily solved. In particular, as the Born–Oppenheimer representation is abandoned, (i) a uniquely-defined expression for the classical force driving nuclear dynamics in non-adiabatic conditions naturally emerges, while (ii) the concept of conical intersections disappears. These features of the exact factorization suggest promising applications in the field of ab initio molecular dynamics. The purpose of this chapter is to justify this statement and thoroughly discuss the properties of the exact factorization.

The theoretical framework will be introduced in Section 17.2, and will be compared to the Born–Oppenheimer representation in Section 17.3. Here, the theoretical analysis will be supported by numerical data. A strategy yielding a quantum-classical algorithm for ab initio molecular dynamics based on the exact factorization will be described in Section 17.4, along with its application to the photo-induced dynamics of oxirane. In Section 17.5, we briefly present some observations on the molecular Berry phase, by comparing the traditional geometric phase coming from the Born–Oppenheimer approximation with the phase associated with the exact factorization. Conclusions are presented in Section 17.6.

17.2 The Time-Dependent Molecular Problem in the Exact-Factorization Formulation

A system of interacting electrons and nuclei is described, in the non-relativistic limit, by the molecular Hamiltonian

$$\hat{H}(\boldsymbol{r},\boldsymbol{R}) = \sum_{K=1}^{N_n} \frac{-\hbar^2}{2M_K}\nabla_K^2 + \hat{T}_e(\boldsymbol{r}) + \hat{V}_{ee}(\boldsymbol{r}) + \hat{V}_{nn}(\boldsymbol{R}) + \hat{V}_{en}(\boldsymbol{r},\boldsymbol{R}) \tag{17.1}$$

$$= \sum_{K=1}^{N_n} \frac{-\hbar^2}{2M_K}\nabla_K^2 + \hat{H}_{BO}(\mathbf{r},\mathbf{R}). \tag{17.2}$$

The first term on the right-hand side is the nuclear kinetic energy operator, with ∇_K the gradient with respect to the position of the Kth nucleus, and M_K its mass; $\hbar$ is the reduced Planck's constant. The symbols $\boldsymbol{r}$ and $\boldsymbol{R}$ are used to indicate the sets of $3N_{el}$ and $3N_n$ electronic and nuclear coordinates, respectively. The electronic kinetic energy operator is indicated as $\hat{T}_e(\boldsymbol{r})$, whereas the electron–electron, nucleus–nucleus, and electron–nucleus interaction potentials are indicated as $\hat{V}_{ee}(\boldsymbol{r})$, $\hat{V}_{nn}(\boldsymbol{R})$, and $\hat{V}_{en}(\boldsymbol{r},\boldsymbol{R})$, respectively. The sum of the electronic kinetic energy and of the interaction potentials is denoted $\hat{H}_{BO}(\boldsymbol{r},\boldsymbol{R})$, i.e., the Born–Oppenheimer (BO) Hamiltonian. In what follows, we will not consider explicitly time-dependent Hamiltonians, where, for instance, a classical time-dependent field acts on the electrons, $\hat{V}_{ext}(\boldsymbol{r},t)$, and/or on the nuclei, $\hat{V}_{ext}(\boldsymbol{R},t)$. However, such contribution can be easily included.

17.2.1 Wave Function Ansatz

The time-dependent Schrödinger equation (TDSE)

$$\hat{H}(\boldsymbol{r},\boldsymbol{R})\Psi(\boldsymbol{r},\boldsymbol{R},t) = i\hbar\frac{\partial}{\partial t}\Psi(\boldsymbol{r},\boldsymbol{R},t), \tag{17.3}$$

determines the evolution of the molecular wave function, $\Psi(\boldsymbol{r},\boldsymbol{R},t)$. As shown in Refs. [35, 36], the molecular wave function can be factored, without making any approximation, as

$$\Psi(\boldsymbol{r},\boldsymbol{R},t) = \Xi(\boldsymbol{R},t)\Phi_{\boldsymbol{R}}(\boldsymbol{r},t), \tag{17.4}$$

with $\Xi(\boldsymbol{R},t)$ the nuclear wave function, and $\Phi_{\boldsymbol{R}}(\boldsymbol{r},t)$ an electronic factor that depends parametrically on the nuclear positions $\boldsymbol{R}$. Existence and uniqueness, up to a gauge transformation, can be easily proven by imposing the partial normalization condition

$$\int d\boldsymbol{r}|\Phi_{\boldsymbol{R}}(\boldsymbol{r},t)|^2 = 1 \quad \forall\, \boldsymbol{R},t \tag{17.5}$$

on the electronic wave function. It follows from Eq. (17.5) that the exact nuclear density is

$$\int d\boldsymbol{r}|\Psi(\boldsymbol{r},\boldsymbol{R},t)|^2 = |\Xi(\boldsymbol{R},t)|^2. \tag{17.6}$$

Therefore, the nuclear wave function can be written as

$$\Xi(\boldsymbol{R},t) = \exp\left[\frac{i}{\hbar}S(\boldsymbol{R},t)\right]\sqrt{\int d\boldsymbol{r}|\Psi(\boldsymbol{r},\boldsymbol{R},t)|^2}. \tag{17.7}$$

Here, $S(\boldsymbol{R},t)$ is a real function of nuclear positions and time, which will be discussed below. The electronic wave function can be determined inverting Eq. (17.4)

$$\Phi_{\boldsymbol{R}}(\boldsymbol{r},t) = \frac{\Psi(\boldsymbol{r},\boldsymbol{R},t)}{\exp\left[\frac{i}{\hbar}S(\boldsymbol{R},t)\right]|\Xi(\boldsymbol{R},t)|}. \tag{17.8}$$

Equation (17.7) and (17.8) prove the existence of the factored form of the molecular wave function: they are proof that a molecular wave function can be rewritten as the single product given in Eq. (17.4). The electronic $\Phi_{\boldsymbol{R}}(\boldsymbol{r},t)$ and nuclear $\Xi(\boldsymbol{R},t)$ wave functions can be used to construct the total wave function $\Psi(\boldsymbol{r},\boldsymbol{R},t)$ at all times, provided that they evolve according to the equations of motion derived in Section 17.2.2. If, for a particular nuclear configuration $\boldsymbol{R}$, the nuclear density is zero (thus $\Xi(\boldsymbol{R},t)$ has a node), $|\Psi(\boldsymbol{r},\boldsymbol{R},t)|^2$ has to be zero as well, $\forall\boldsymbol{r}$. In fact, in Eq. (17.6) the nuclear density is given as the sum of positive-definite terms. In this case, the electronic wave function from Eq. (17.8) might still be well-defined, being the ratio of two quantities both tending to zero.

The factored form of the molecular wave function in Eq. (17.4) is invariant under a gauge-like transformation of the electronic and nuclear wave functions: if $\Xi(\boldsymbol{R},t)$ is multiplied by a phase factor that only depends on $\boldsymbol{R}$ and t, and $\Phi_{\boldsymbol{R}}(\boldsymbol{r},t)$ is multiplied by its complex conjugate, the final product remains unchanged. We prove below that this gauge freedom is the only freedom in the definition of $\Xi(\boldsymbol{R},t)$ and $\Phi_{\boldsymbol{R}}(\boldsymbol{r},t)$ that shall be dealt with by making a *choice of gauge*. Suppose that two different products yield the same total wave function, namely

$$\Psi(\boldsymbol{r},\boldsymbol{R},t) = \Xi(\boldsymbol{R},t)\Phi_{\boldsymbol{R}}(\boldsymbol{r},t) = \widetilde{\Xi}(\boldsymbol{R},t)\widetilde{\Phi}_{\boldsymbol{R}}(\boldsymbol{r},t). \tag{17.9}$$

We can divide the second and third term by $\Xi(\boldsymbol{R},t)$ and by $\tilde{\Phi}_{\boldsymbol{R}}(\boldsymbol{r},t)$, obtaining

$$\frac{\widetilde{\Xi}(\boldsymbol{R},t)}{\Xi(\boldsymbol{R},t)} = \frac{\Phi_{\boldsymbol{R}}(\boldsymbol{r},t)}{\widetilde{\Phi}_{\boldsymbol{R}}(\boldsymbol{r},t)} = g(\boldsymbol{R},t). \tag{17.10}$$

The quantity on the left-hand side depends on $(\boldsymbol{R},t)$ only, and we have indicated it as $g(\boldsymbol{R},t)$. Equation (17.10) provides a relation between the electronic wave functions $\Phi_{\boldsymbol{R}}(\boldsymbol{r},t)$ and $\widetilde{\Phi}_{\boldsymbol{R}}(\boldsymbol{r},t)$, namely

$$\Phi_{\boldsymbol{R}}(\boldsymbol{r},t) = g(\boldsymbol{R},t)\widetilde{\Phi}_{\boldsymbol{R}}(\boldsymbol{r},t). \tag{17.11}$$

For both $\Phi_{\boldsymbol{R}}(\boldsymbol{r},t)$ and $\widetilde{\Phi}_{\boldsymbol{R}}(\boldsymbol{r},t)$, the partial normalization condition has to hold. By imposing such a condition in both sides of the equality,

$$1 = \int d\boldsymbol{r}|\Phi_{\boldsymbol{R}}(\boldsymbol{r},t)|^2 = |g(\boldsymbol{R},t)|^2\int d\boldsymbol{r}\left|\widetilde{\Phi}_{\boldsymbol{R}}(\boldsymbol{r},t)\right|^2 = |g(\boldsymbol{R},t)|^2, \tag{17.12}$$

we find that the only freedom in the definition of $\Xi(\boldsymbol{R},t)$ and $\Phi_{\boldsymbol{R}}(\boldsymbol{r},t)$ lies in a phase factor $g(\boldsymbol{R},t) = \exp[(i/\hbar)\theta(\boldsymbol{R},t)]$, where $\theta(\boldsymbol{R},t)$ is a real function of nuclear positions and time. Clearly, the total

molecular wave function is invariant under such $(\boldsymbol{R}, t)$-dependent gauge transformations,

$$\Xi(\boldsymbol{R}, t) \rightarrow \widetilde{\Xi}(\boldsymbol{R}, t) = \exp\left[-\frac{i}{\hbar}\theta(\boldsymbol{R}, t)\right]\Xi(\boldsymbol{R}, t) \tag{17.13}$$

$$\Phi_{\boldsymbol{R}}(\boldsymbol{r}, t) \rightarrow \widetilde{\Phi}_{\boldsymbol{R}}(\boldsymbol{r}, t) = \exp\left[\frac{i}{\hbar}\theta(\boldsymbol{R}, t)\right]\Phi_{\boldsymbol{R}}(\boldsymbol{r}, t). \tag{17.14}$$

Uniqueness of Eq. (17.4) is guaranteed upon fixing the gauge freedom with a suitable choice of $\theta(\boldsymbol{R}, t)$. The gauge function $\theta(\boldsymbol{R}, t)$ can be absorbed into the phase $S(\boldsymbol{R}, t)$ of the nuclear wave function, since choosing $S(\boldsymbol{R}, t)$ is fully equivalent to choosing the gauge. This point will be clarified in the sections below, where we discuss possible ways of fixing the gauge freedom.

17.2.2 Equations of Motion

As illustrated in great detail in Ref. [38], inserting the factored form of the molecular wave function, Eq. (17.4), in the TDSE (17.3), and using the partial normalization condition (17.5) [39, 40], the evolution equations

$$\left[\sum_{K=1}^{N_n}\frac{[-i\hbar\nabla_K + \boldsymbol{A}_K(\boldsymbol{R}, t)]^2}{2M_K} + \epsilon(\boldsymbol{R}, t)\right]\Xi(\boldsymbol{R}, t) = i\hbar\frac{\partial}{\partial t}\Xi(\boldsymbol{R}, t) \tag{17.15}$$

$$[\hat{H}_{BO}(\boldsymbol{r}, \boldsymbol{R}) + \hat{U}_{en}[\Phi_{\boldsymbol{R}}, \Xi] - \epsilon(\boldsymbol{R}, t)]\Phi_{\boldsymbol{R}}(\boldsymbol{r}, t) = i\hbar\frac{\partial}{\partial t}\Phi_{\boldsymbol{R}}(\boldsymbol{r}, t), \tag{17.16}$$

can be derived. Equation (17.15), the nuclear equation, has the form of a standard TDSE, where a time-dependent vector potential and a time-dependent scalar potential represent the effect of the electrons on nuclear dynamics [41–46]. It is worth mentioning here that the nuclear Hamiltonian (the quantity in square brackets in Eq. (17.15)), and in particular the time-dependent scalar potential $\epsilon(\boldsymbol{R}, t)$), does not depend on Ξ. We clarify this statement in the next paragraph where we introduce its explicit expression. Since no approximation has been invoked so far in the derivation of nuclear (17.15) and electronic (17.16) equations, those potentials can incorporate excited-state effects. One can show that $\Xi(\boldsymbol{R}, t)$ is a genuine nuclear wave function [38], because (i) it evolves according to the TDSE (17.15), (ii) yields by definition the exact nuclear N-body density (Eq. (17.6)), and (iii) gives the exact nuclear N-body current density,

$$\mathbf{J}_K(\boldsymbol{R}, t) = \frac{1}{M_K}[\mathrm{Im}(\Xi^*(\boldsymbol{R}, t)\hbar\nabla_K\Xi(\boldsymbol{R}, t)) + |\Xi(\boldsymbol{R}, t)|^2\mathbf{A}_K(\boldsymbol{R}, t)]. \tag{17.17}$$

In analogy to the expression of the nuclear density given in Eq. (17.6), Eq. (17.17) can be derived from its definition in terms of the full wave function, i.e., $\boldsymbol{J}_K(\boldsymbol{R}, t) = [\hbar/M_K]\int d\boldsymbol{r}\mathrm{Im}[\Psi^*(\boldsymbol{r}, \boldsymbol{R}, t)\nabla_K\Psi(\boldsymbol{r}, \boldsymbol{R}, t)]$, introducing the factored form of $\Psi(\boldsymbol{r}, \boldsymbol{R}, t)$.

In Eqs. (17.15) and (17.16) the terms responsible for the dynamical, i.e., beyond the adiabatic BO approximation, coupling between electronic and nuclear motion are the time-dependent vector potential $\mathbf{A}_\nu(\boldsymbol{R}, t)$ [46, 47],

$$\boldsymbol{A}_K(\boldsymbol{R}, t) = \langle\Phi_{\boldsymbol{R}}(t)| - i\hbar\nabla_K\Phi_{\boldsymbol{R}}(t)\rangle_{\boldsymbol{r}}, \tag{17.18}$$

the electron–nuclear coupling operator $\hat{U}_{en}[\Phi_{\boldsymbol{R}}, \Xi]$ [35, 36, 43, 44, 48–52]

$$\hat{U}_{en}[\Phi_{\boldsymbol{R}}, \Xi] = \sum_{K=1}^{N_n}\frac{1}{M_K}\left[\frac{[-i\hbar\nabla_K + \boldsymbol{A}_K(\boldsymbol{R}, t)]^2}{2} + \left(\frac{-i\hbar\nabla_K\Xi(R, t)}{\Xi(\boldsymbol{R}, t)} + \boldsymbol{A}_\nu(\boldsymbol{R}, t)\right)\cdot(-i\hbar\nabla_K - \boldsymbol{A}_K(\boldsymbol{R}, t))\right] \tag{17.19}$$

and the time-dependent potential energy surface (TDPES) $\epsilon(\boldsymbol{R}, t)$ [35, 38, 41, 42, 44–46, 53, 54],

$$\epsilon(\boldsymbol{R}, t) = \langle \Phi_{\boldsymbol{R}}(t) | \hat{H}_{BO} + \hat{U}_{en} - i\hbar \frac{\partial}{\partial t} | \Phi_{\boldsymbol{R}}(t) \rangle_{\boldsymbol{r}}. \tag{17.20}$$

The symbol $\langle \cdot \rangle_{\boldsymbol{r}}$ indicates an integration over electronic coordinates only. The TDPES is expressed in Eq. (17.20) as the "expectation value" on $\Phi_{\boldsymbol{R}}(\boldsymbol{r}, t)$ of the sum of three operators: $\hat{H}_{BO}$, $\hat{U}_{en}$, and $-i\hbar \frac{\partial}{\partial t}$. The term depending on the electron–nuclear coupling operator seems to depend on $\Xi(\boldsymbol{R}, t)$ via the second term in square brackets in Eq. (17.19). However, $\langle \Phi_{\boldsymbol{R}}(t)| - i\hbar \nabla_K - \boldsymbol{A}_K(\boldsymbol{R}, t) | \Phi_{\boldsymbol{R}}(t) \rangle_{\boldsymbol{r}} = \langle \Phi_{\boldsymbol{R}}(t)| - i\hbar \nabla_K | \Phi_{\boldsymbol{R}}(t) \rangle_{\boldsymbol{r}} - \boldsymbol{A}_K(\boldsymbol{R}, t) = 0$, which follows from the definition of the time-dependent vector potential of Eq. (17.18) and the partial normalization condition. Therefore, $\epsilon(\boldsymbol{R}, t)$ and the nuclear Hamiltonian in Eq. (17.15) do not depend on Ξ.

Under the gauge transformations (17.13) and (17.14), the scalar potential and the vector potential transform as

$$\tilde{\epsilon}(\boldsymbol{R}, t) = \epsilon(\boldsymbol{R}, t) + \frac{\partial}{\partial t} \theta(\boldsymbol{R}, t) \tag{17.21}$$

$$\tilde{\boldsymbol{A}}_K(\boldsymbol{R}, t) = \mathbf{A}_K(\boldsymbol{R}, t) + \nabla_K \theta(\boldsymbol{R}, t) \, . \tag{17.22}$$

The evolution equations, (17.15) and (17.16), are form-invariant under such a transformation.

A large amount of literature can be found on different topics related to the exact factorization of the electron–nuclear wave function. For instance, still in the time-dependent context, studies have focused on proposing trajectory-based solutions [55–63] of the coupled nuclear (17.15) and electronic (17.16) equations, on introducing a perturbation-theory framework to treat situations of weak non-adiabatic coupling between electronic and nuclear motion [48–51] and to cure the inconsistencies of the BO approximation, on deriving a density-functional theory of the coupled electron–nuclear problem [64–66], on exploiting the "inverse" exact-factorization formalism to derive the exact electronic TDSE with non-classical nuclei [67–69], on reformulating the dynamical problem for systems different from electrons and nuclei, as purely electronic systems [70] or electron-photon systems [71, 72]. The static formulation of the exact factorization, relying on the time-independent Schrödinger equation, has also been the subject of extensive studies that focus on the properties of the wave functions and of the potentials [73–85], or, interestingly, on tackling the problem of geometric phases in molecular problems [86–88].

A review of all these topics is beyond the scope of this chapter. Therefore, we will present here (i) the relationship between the exact factorization and the, more standard, BO representation of the electron–nuclear problem, based on the analysis of the time-dependent potentials of the theory, (ii) a procedure to solve – approximately – the coupled electronic and nuclear equations, (iii) the exact-factorization perspective on molecular geometric phases.

17.3 The Born–Oppenheimer Framework and the Exact Factorization

Standard approaches to formulate and solve the electron–nuclear dynamical problem rely on the Born–Huang expansion of the time-dependent molecular wave function,

$$\Psi(\boldsymbol{r}, \boldsymbol{R}, t) = \sum_{\alpha} \Xi^{(\alpha)}_{BO}(\boldsymbol{R}, t) \Phi^{(\alpha)}_{\boldsymbol{R}}(\boldsymbol{r}), \tag{17.23}$$

which we have previously also referred to as the BO framework. Here, $\{\Phi^{(\alpha)}_{\boldsymbol{R}}(\boldsymbol{r})\}_{\alpha=1,\dots}$ are the eigenstates of the (electronic) BO Hamiltonian $\hat{H}_{BO}$, determined for each value of the parameter $\boldsymbol{R}$, and $\Xi^{(\alpha)}_{BO}(\boldsymbol{R}, t)$ are expansion coefficients. The corresponding eigenvalues are denoted $\epsilon^{(\alpha)}_{BO}(\boldsymbol{R})$, the adiabatic, or BO, potential energy surfaces (PESs). We introduce also the matrix of non-adiabatic

coupling vectors, $\boldsymbol{d}_{\beta\alpha,K}(\boldsymbol{R}) = \langle \Phi_{\boldsymbol{R}}^{(\beta)} | \nabla_K \Phi_{\boldsymbol{R}}^{(\alpha)} \rangle_{\boldsymbol{r}}$, which are three-dimensional vectors for each value of the nuclear index K, to be used in the following. When only one term of the expansion (17.23) is retained, the resulting molecular wave function is known as the *adiabatic approximation*. Since the adiabatic states are orthonormal, the nuclear density following from Eq. (17.23) is

$$|\Xi(\boldsymbol{R},t)|^2 = \sum_\alpha |\Xi_{BO}^{(\alpha)}(\boldsymbol{R},t)|^2. \tag{17.24}$$

The quantity $\Xi_{BO}^{(\alpha)}(\boldsymbol{R},t)$ is sometimes indicated as the adiabatic contribution of the nuclear wave packet in state α. Being a proper electronic basis, the adiabatic states can be used to expand the electronic wave function of the exact factorization (17.4) as well,

$$\Phi_{\boldsymbol{R}}(\boldsymbol{r},t) = \sum_\alpha C_\alpha(\boldsymbol{R},t)\Phi_{\boldsymbol{R}}^{(\alpha)}(\boldsymbol{r}), \tag{17.25}$$

where the coefficients satisfy the relation

$$\sum_\alpha |C_\alpha(\boldsymbol{R},t)|^2 = 1 \quad \forall\, \boldsymbol{R},t \tag{17.26}$$

by virtue of the partial normalization condition (17.5). It is interesting to note that, even though the nuclear wave packet might have contributions, i.e., the $\Xi_{BO}^{(\alpha)}(\boldsymbol{R},t)$, in different electronic states, the single-product representation of the molecular wave function still holds. In fact,

$$\Psi(\boldsymbol{r},\boldsymbol{R},t) = \left(e^{\frac{i}{\hbar}S(\boldsymbol{R},t)} \sqrt{\sum_\alpha |\Xi_{BO}^{(\alpha)}(\boldsymbol{R},t)|^2} \right) \left(\sum_\alpha C_\alpha(\boldsymbol{R},t)\Phi_{\boldsymbol{R}}^{(\alpha)}(\boldsymbol{r}) \right) \tag{17.27}$$

which follows from Eqs. (17.7) and (17.25), where the first term in parenthesis is the nuclear wave function $\Xi(\boldsymbol{R},t)$ and the second term is the electronic wave function $\Phi_{\boldsymbol{R}}(\boldsymbol{r},t)$.

The identity relation between Eqs. (17.4) and (17.23),

$$\Xi(\boldsymbol{R},t)\sum_\alpha C_\alpha(\boldsymbol{R},t)\Phi_{\boldsymbol{R}}^{(\alpha)}(\boldsymbol{r}) = \sum_\alpha \Xi_{BO}^{(\alpha)}(\boldsymbol{R},t)\Phi_{\boldsymbol{R}}^{(\alpha)}(\boldsymbol{r}) \tag{17.28}$$

yields a relationship between the expansion coefficients

$$\Xi(\boldsymbol{R},t)C_\alpha(\boldsymbol{R},t) = \Xi_{BO}^{(\alpha)}(\boldsymbol{R},t). \tag{17.29}$$

To help the reader relate the exact factorization to the standard BO framework, we give below the expressions of the time-dependent vector potential and of the TDPES in terms of adiabatic quantities. Inserting Eq. (17.25) into Eq. (17.18), one obtains

$$\boldsymbol{A}_K(\boldsymbol{R},t) = -i\hbar \left[\sum_\alpha C_\alpha^*(\boldsymbol{R},t)\nabla_K C_\alpha(\boldsymbol{R},t) + \sum_{\alpha,\beta} C_\alpha^*(\boldsymbol{R},t)C_\beta(\boldsymbol{R},t)\boldsymbol{d}_{\alpha\beta,K}(\boldsymbol{R}) \right]. \tag{17.30}$$

When this expression is used in the definition of the electron–nuclear coupling operator, one also gets $\hat{U}_{en}$ in terms of adiabatic quantities. The TDPES in Eq. (17.20) is the sum of three terms, each expressed in the adiabatic basis as

$$\langle \Phi_{\boldsymbol{R}}(t)|\hat{H}_{BO}|\Phi_{\boldsymbol{R}}(t)\rangle_{\boldsymbol{r}} = \sum_\alpha |C_\alpha(\boldsymbol{R},t)|^2 \epsilon_{BO}^{(\alpha)}(\boldsymbol{R}), \tag{17.31}$$

$$\begin{aligned}\langle \Phi_{\boldsymbol{R}}(t)|\hat{U}_{en}|\Phi_{\boldsymbol{R}}(t)\rangle_{\boldsymbol{r}} = &\sum_K \frac{\hbar^2}{M_K} \sum_{\alpha,\beta} \Big[\left|\nabla_K C_\alpha(\boldsymbol{R},t)\right|^2 \delta_{\alpha\beta} + \Big(\left(\nabla_K C_\alpha^*(\boldsymbol{R},t)\right) C_\beta(\boldsymbol{R},t) \\ &- C_\alpha^*(\boldsymbol{R},t)\left(\nabla_K C_\beta(\boldsymbol{R},t)\right) \Big) \cdot \boldsymbol{d}_{\alpha\beta,K}(\boldsymbol{R}) + C_\alpha^*(\boldsymbol{R},t)C_\beta(\boldsymbol{R},t)D_{\alpha\beta,K}(\boldsymbol{R}) \Big] \\ &- \sum_K \frac{\boldsymbol{A}_K^2(\boldsymbol{R},t)}{2M_K},\end{aligned} \tag{17.32}$$

$$\langle \Phi_{\boldsymbol{R}}(t)| - i\hbar\frac{\partial}{\partial t}|\Phi_{\boldsymbol{R}}(t)\rangle_{\boldsymbol{r}} = -i\hbar \sum_\alpha C_\alpha^*(\boldsymbol{R},t)\frac{\partial}{\partial t}C_\alpha(\boldsymbol{R},t). \tag{17.33}$$

The scalar quantity $D_{\alpha\beta,K}(\boldsymbol{R}) = \langle \nabla_K \Phi_{\boldsymbol{R}}^{(\alpha)} | \nabla_K \Phi_{\boldsymbol{R}}^{(\beta)} \rangle_{\boldsymbol{r}}$ in Eq. (17.32) indicates the second-order non-adiabatic coupling.

Equations (17.24)–(17.33) show the relationship between the BO representation of the electron–nuclear wave function and the exact factorization. The key difference between the two representations is that in the latter, the electronic wave function acquires an explicit time dependence. In contrast, in the BO framework, the electrons appear in the problem as the eigenstates of the electronic BO Hamiltonian, providing the static potentials driving nuclear dynamics. Therefore, within a BO perspective of non-adiabatic dynamics, processes are visualized employing static adiabatic PESs that are coupled in regions of nuclear space where they are close or degenerate, i.e., avoided crossings (see Section 17.3.1) or conical intersections (see Section 17.3.2). The nuclear wave packet has adiabatic contributions evolving "on" the different surfaces, and transfers amplitude to other states in the regions of coupling. Such a perspective is revisited within the exact factorization.

In Section 17.3.1, for a one-dimensional model system, the Shin–Metiu model [89], we will show that a single PES, uniquely defined up to a gauge [35, 36, 41, 42, 44, 53, 54], that changes with time is able to guide the nuclear dynamics in different ways in different portions of nuclear configuration space. It is able to achieve this result simply by changing its shape over time: it has a diabatic shape in the region where the BO PESs are close in energy [41, 42, 44]; it develops *steps* connecting adiabatic shapes, forcing the wave packet to spatially split onto BO PESs of different slopes [41, 42, 44]; it has a mean-field character when the adiabatic contributions of a nuclear wave packet are localized in the same region [45]; it presents oscillations to account for interference effects [45]. Clearly, the TDPES encodes all features of (electronic) dynamics, and this information is transferred to the nuclear TDSE.

In Section 17.3.2, we will generalize our observations on the TDPES of Section 17.3.1 to a two-dimensional case. Additionally, we will discuss the time-dependent vector potential. The vector potential couples to the momentum in the nuclear TDSE [44], and it can be related to the nuclear velocity field [46, 47], as shown in Eq. (17.17). The feature highlighted in this section is that, in general, the time-dependent vector potential is not irrotational, meaning that it cannot be written as the gradient of a scalar function of $\boldsymbol{R}$ and t. The consequence is that the vector potential cannot be gauged away in general situations, a property that is important to keep in mind when approximations to the exact evolution equations, Eqs. (17.16) and (17.15), are to be developed.

17.3.1 One-Dimensional Case: Time-Dependent Potential Energy Surface

To proceed with the analysis of the TDPES and its relation to the static BO PESs, it is instructive to decompose the TDPES into gauge-invariant (GI) and gauge-dependent (GD) components, $\epsilon(\boldsymbol{R}, t) = \epsilon_{GI}(\boldsymbol{R}, t) + \epsilon_{GD}(\boldsymbol{R}, t)$, where

$$\epsilon_{GI}(\boldsymbol{R}, t) = \langle \Phi_{\boldsymbol{R}}(t) | \hat{H}_{BO} | \Phi_{\boldsymbol{R}}(t) \rangle_{\boldsymbol{r}} + \sum_{K=1}^{N_n} \frac{\hbar^2}{2M_K} \langle \nabla_K \Phi_{\boldsymbol{R}}(t) | \nabla_K \Phi_{\boldsymbol{R}}(t) \rangle_{\boldsymbol{r}} - \sum_{\nu=1}^{N_n} \frac{\boldsymbol{A}_K^2(\boldsymbol{R}, t)}{2M_K} \tag{17.34}$$

and

$$\epsilon_{GD}(\boldsymbol{R}, t) = \langle \Phi_{\boldsymbol{R}}(t) | - i\hbar \frac{\partial}{\partial t} | \Phi_{\boldsymbol{R}}(t) \rangle_{\boldsymbol{r}}. \tag{17.35}$$

The second and third terms on the right-hand side of Eq. (17.34) are obtained from the action of the electron–nuclear coupling operator of Eq. (17.19) on the electronic wave function. The GI part of the TDPES, ϵ_{GI}, is invariant under the gauge transformation (17.14): $\tilde{\epsilon}_{GI}(\boldsymbol{R}, t) = \epsilon_{GI}(\boldsymbol{R}, t)$; the GD part, on the other hand, transforms as $\tilde{\epsilon}_{GD}(\boldsymbol{R}, t) = \epsilon_{GD}(\boldsymbol{R}, t) + \frac{\partial}{\partial t}\theta(\boldsymbol{R}, t)$.

The analysis of the TDPES will be based, in this section, on a one-dimensional model for proton-coupled electron transfer [89], the details of which are given in Refs. [41–44]. The model consists of two positively-charged ions fixed at a distance L, a positively-charged ion of mass M (in this case, the proton mass is used) moving in one dimension between the two fixed ions, thus interacting with the fixed ions via a bare Coulomb potential, an electron moving in one dimension and interacting with all ions via a soft-Coulomb potential. The schematic representation of the model is given in Figure 17.1.

Since we are dealing with a one-dimensional problem, we can choose a gauge such that the time-dependent vector potential is always zero. Inverting Eq. (17.4) to express $\Phi_{\boldsymbol{R}}(\boldsymbol{r},t)$ in terms of $\Psi(\boldsymbol{r},\boldsymbol{R},t)$ and of $\Xi(\boldsymbol{R},t)$, and inserting such an expression in the definition of the vector potential (17.18), we obtain a relationship between the vector potential itself, the nuclear velocity field, and the phase $S(\boldsymbol{R},t)$ of the nuclear wave function

$$\boldsymbol{A}_K(\boldsymbol{R},t) = \frac{\mathrm{Im}\langle\Psi(t)|\hbar\nabla_K\Psi(t)\rangle_{\boldsymbol{r}}}{|\Xi(\boldsymbol{R},t)|^2} - \nabla_K S(\boldsymbol{R},t). \tag{17.36}$$

Adopting a one-dimensional representation of the vectors, $A(R,t) = \hbar\mathrm{Im}\langle\Psi(t)|\frac{\partial}{\partial R}\Psi(t)\rangle_r/|\Xi(R,t)|^2 - \frac{\partial}{\partial R}S(R,t)$, and imposing $A(R,t)=0$, leads to an expression that defines the phase of the nuclear wave function,

$$S(R,t) = \int^R dR' \frac{\hbar\mathrm{Im}\langle\Psi(t)|\frac{\partial}{\partial R'}\Psi(t)\rangle_r}{|\Xi(R',t)|^2}\,, \tag{17.37}$$

in this gauge. Since the time-dependent vector potential is identically zero all along the dynamics, only the TDPES appears in the nuclear TDSE (17.15). Therefore, the TDPES is the only potential affecting the nuclear dynamics and carrying information about electronic dynamics.

The dynamics is initiated by setting a Gaussian wave packet, with zero average velocity, on the first excited adiabatic state S_1, centered at $R_0 = -4$ bohr and with variance $\sigma = 1/\sqrt{2.85}$ bohr. Figure 17.1 shows the BO PESs corresponding to the ground S_0 and first-excited S_1 states (which are the only states populated during the dynamics). The time evolution of the molecular wave function is generated by solving the full TDSE by using the split-operator technique [90] with time step $dt = 2.4\ 10^{-3}$ fs (0.1 a.u.). Therefore, in this analysis, the exact-factorization equations (17.15) and (17.16) are not directly solved to determine the nuclear $\Xi(\boldsymbol{R},t)$ and electronic $\Phi_{\boldsymbol{R}}(\boldsymbol{r},t)$ wave functions. Instead, using Eq. (17.7) and fixing the gauge as in Eq. (17.37), $\Xi(\boldsymbol{R},t)$ is computed from $\Psi(\boldsymbol{r},\boldsymbol{R},t)$ and $\Phi_{\boldsymbol{R}}(\boldsymbol{r},t)$ from Eq. (17.8). A numerical study on the direct integration of Eqs. (17.15)

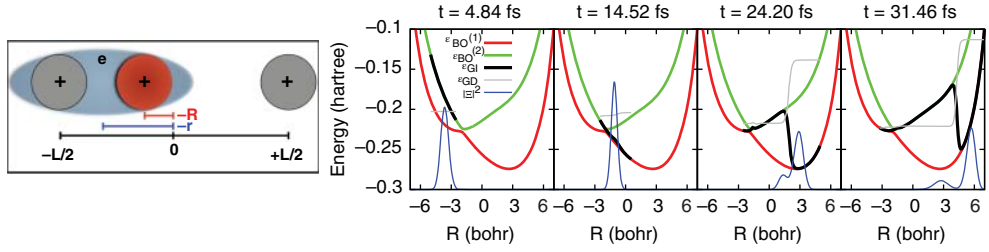


Figure 17.1 Left: Schematic representation of the model for proton coupled electron transfer. Right: GI part (black) and GD part (gray) of the TDPES at times $t = 4.84, 14.52, 24.20, 31.46$ fs. The two lowest BO surfaces are shown for reference. The TDPES is shown in the regions where the nuclear density (blue lines in the panels) is larger than 10^{-8}. The GD part of the TDPES is shifted in all panels to superimpose it on the plot of the GI part.

and (17.16) can be found in Ref. [91]. Once the electronic wave function is known, Eqs. (17.34) and (17.35) allow one to determine the TDPES. We will propose here a qualitative analysis of the TDPES and of its relation to the BO PESs. A more quantitative discussion can be found in Refs. [41–45].

Figure 17.1 shows different snapshots along the dynamics of the nuclear density (blue curves) and of the TDPES, decomposed in its GI (black lines) and GD (gray lines) components. The nuclear dynamics is initiated in the first excited state, corresponding to the green curve, in the negative R-region. Then, the nuclear wave packet slides down towards the avoided crossing, where the energy gap between S_0 (corresponding to the red curve) and S_1 closes. While crossing the region of coupling, amplitude is transferred to the ground state, producing a two-component wave packet evolving with different speeds depending on the shapes of the corresponding BO PES: in the positive R-region the red curve accelerates the wave packet, whereas the green curve slows it down.

At early times ($t = 4.84$ fs), the dynamics is driven simply by the excited state, as the GI part of the TDPES completely lies on the green curve; later on ($t = 14.52$ fs), ϵ_{GI} has a diabatic shape, connecting the adiabatic surfaces smoothly through the avoided crossing, thus allowing amplitude to "flow" from S_1 to S_0; finally ($t = 24.20$ fs and $t = 31.46$ fs), the GI component of the TDPES presents different adiabatic shapes in different portions of R-space connected by *steps*, clearly suggesting that the nuclear wave packet is driven by different forces on the two sides of the steps. The spatial splitting of the nuclear density is, therefore, attributable to the appearance of the steps. However, one should keep in mind that the GD component of the TDPES also has an effect on the dynamics. As shown in Figure 17.1, ϵ_{GD} is either constant or piecewise constant, with steps appearing at the same positions as the steps in ϵ_{GI}. Therefore, ϵ_{GD} does not modify the slope of the GI component, it only reduces the size of the steps. These observations have been supported by an analytical justification in Refs. [41, 44], which proves their general validity. In particular, in Ref. [44] it is demonstrated that the height of the step in ϵ_{GD} in the region where the nuclear density splits (for instance, at around $R = 4$ bohr in the last panel at $t = 31.46$ fs of Figure 17.1) approximately amounts to the energy difference between the BO PESs, similar to the height of the step in ϵ_{GI}.

According to Eq. (17.24), the nuclear density can be decomposed into adiabatic contributions, and this property is verified in Figure 17.2. Initially at time $t = 4.84$ fs, the nuclear density has only the contribution corresponding to S_1 (green curve). When the wave packet crosses the non-adiabatic coupling region at time $t = 14.52$ fs, the ground-state and excited-state contributions are localized in the same region. The spatial splitting appears at later times, $t = 24.20$ fs and

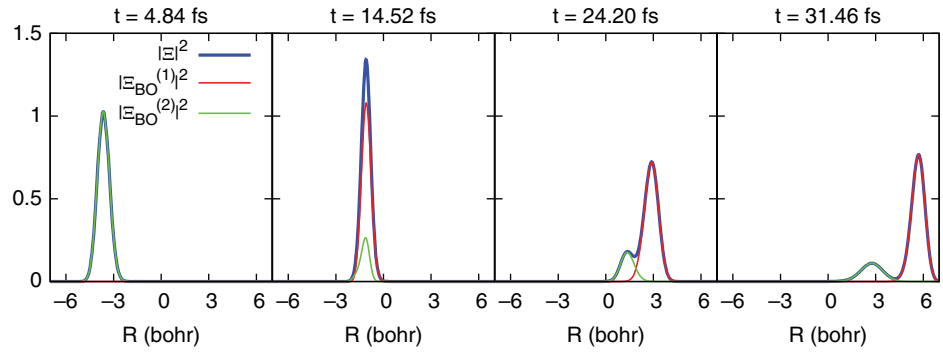


Figure 17.2 Nuclear density $|\Xi(R,t)|^2$ (blue lines), along with its BO contributions $|\Xi^{(1)}_{BO}(R,t)|^2$ (red lines) and $|\Xi^{(2)}_{BO}(R,t)|^2$ (green lines) at times $t = 4.84, 14.52, 24.20, 31.46$ fs.

$t = 31.46$ fs, and in Figure 17.2 it is evident that the S_0 component (red curve) is localized in the same region where the GI part of the TDPES has the shape of the lower BO PES, while the S_1 component is localized on the other side of the step in the TDPES.

The coefficients $C_\alpha(R,t)$ for $\alpha = 1, 2$ (corresponding to S_0 and S_1, respectively) of the expansion of the electronic wave function on the adiabatic basis, Eq. (17.25), have a similar behavior, that is analyzed in Figure 17.3 by reporting their squared moduli. The quantity $|C_\alpha(R,t)|^2$ is determined from Eq. (17.29), thus it is only shown in regions where the nuclear density $|\Xi|^2$ is (numerically) larger than zero. In general, $|C_\alpha(R,t)|^2$ is equal to unity if the corresponding portion of the nuclear density is fully localized in one electronic state (see for instance the green curve at time $t = 4.84$ fs or the red curve at $t = 24.20$ fs for $R > 2$ bohr). The plots of the coefficients at the final times of the dynamics present a sigmoid shape, and cross when the nuclear wave packet changes character, i.e., when it splits into different components evolving on different BO surfaces. This observation is fundamental for accounting for the spatial nuclear splitting, in fact, if expansion (17.25) is inserted in the first term on the right-hand side of Eq. (17.34) one gets $\langle \Phi_{\boldsymbol{R}}(t)|\hat{H}_{BO}|\Phi_{\boldsymbol{R}}(t)\rangle_{\boldsymbol{r}} = \sum_\alpha |C_\alpha(\boldsymbol{R},t)|^2\, \epsilon^{(\alpha)}_{BO}(\boldsymbol{R})$. This expression is *simply* a weighted average of the BO PESs, however if this average is computed at time $t = 31.46$ fs, it yields a form of the GI part of the TDPES that is far from a smooth potential, as the steps in $|C_\alpha(R,t)|^2$ translate to the steps in the time-dependent potential [92].

It is instructive at this point to take a look at the time-dependent vector potential, even though in a one-dimensional problem the vector potential can always be gauged away by solving Eq. (17.37). If a function $\theta(R,t)$ is introduced to transform the electronic and nuclear wave functions (as indicated in Eqs. (17.13) and (17.14)) such that $\tilde{\epsilon}_{GD}(R,t) = \epsilon_{GD}(R,t) + \frac{\partial}{\partial t}\theta(R,t) = 0$ in the new gauge, the vector potential becomes non-zero, $\tilde{A}(R,t) = -\int_0^t dt' \epsilon_{GD}(R,t')$, and its action can be related to the momentum-rescaling prescription of surface-hopping algorithms [12, 44]. We analyze below the time-dependent vector potential in this new gauge, and we report numerical results in Figure 17.4 for the same model system as in Figure 17.1.

The time-dependent vector potential appears in the nuclear Hamiltonian of Eq. (17.15) as a contribution to the nuclear momentum. Therefore, in order to investigate its effect, we evaluate the average nuclear momentum, $P_N(t) = \langle \Psi(t)| -i\hbar\frac{\partial}{\partial R}|\Psi(t)\rangle_{r,R} = \int dR[\frac{\partial}{\partial R}S(R,t) + A(R,t)]|\Xi(R,t)|^2$, and analyze the momentum density $\frac{\partial}{\partial R}S(R,t) + A(R,t)$. In Figure 17.4, at different time-steps along the dynamics, we report the vector potential (red lines), the phase contribution to the total momentum density (green lines), i.e., the term $\frac{\partial}{\partial R}S(R,t)$, and their sum (black lines). The effect of the vector potential is basically zero at early times, since the nuclear wave packet is propagating

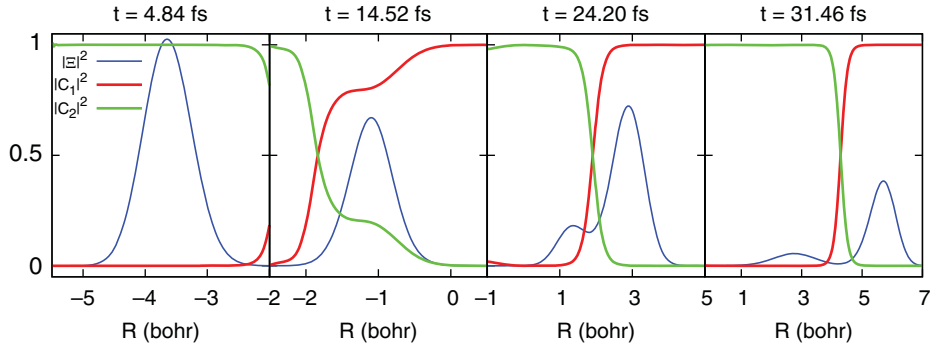


Figure 17.3 Nuclear density $|\Xi(R,t)|^2$ (blue lines), along with the squared moduli of the coefficients $|C_1(R,t)|^2$ (red lines) and $|C_2(R,t)|^2$ (green lines) of the expansion (17.25) at times $t = 4.84, 14.52, 24.20, 31.46$ fs.

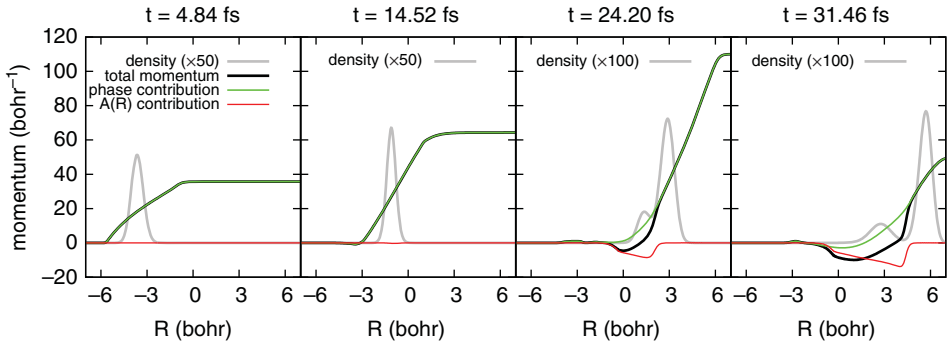


Figure 17.4 Components of the nuclear momentum density at times $t = 4.84, 14.52, 24.20, 31.46$ fs in the gauge $\epsilon_{GD}(R,t) = 0$. In black the full momentum density $\frac{\partial}{\partial R}S(R,t) + A(R,t)$ is shown. The phase contribution, i.e., $\frac{\partial}{\partial R}S(R,t)$, is represented in green, while the vector potential $A(R,t)$ is represented in red. The nuclear density is shown in gray to highlight the region where it is substantially larger than zero (thus where the numerical results are reliable).

under the effect of a single BO PES. After the passage through the avoided crossing located at -2 bohr, it becomes non-zero (it is negative) mainly in the region of R-space corresponding to the portion of the nuclear wave packet propagating "on" S_1. With this choice of gauge, the vector potential acts by slowing down the portion of the wave packet that is higher in energy. Therefore, its action is reminiscent of the momentum corrections applied to the classical trajectories that are propagated according to the surface-hopping algorithm [12, 44]: surface hopping prescribes that, when a trajectory undergoes a non-adiabatic hop from one BO PES to another one that is higher in energy, such a trajectory loses its kinetic energy via a discontinuous rescaling of the momentum.

17.3.2 Two-Dimensional Case: Time-Dependent Potential Energy Surface and Time-Dependent Vector Potential

The purpose of this section is to generalize to higher dimensions the observations presented in Section 17.3.1 concerning the TDPES and to analyze the time-dependent vector potential in a situation where it cannot be set to zero by any choice of gauge. To this end, we will be dealing with a two-dimensional model and a more general choice of gauge than in the previous section will be made. We choose the gauge such that the nuclear wave function $\Xi(\boldsymbol{R},t)$ is real and non-negative, i.e., $\Xi = |\Xi|\ \forall\ \boldsymbol{R},t$, achieved by imposing the condition $S(\boldsymbol{R},t) = 0$. The time-dependent vector potential is a two-dimensional vector field depending on the nuclear coordinates $\boldsymbol{R} = X, Y$ and time t, and with this choice of gauge the expression for the X-component is

$$A_X(\boldsymbol{R},t) = \frac{\mathrm{Im}\langle\Psi(t)|\hbar\frac{\partial}{\partial X}|\Psi(t)\rangle_{\boldsymbol{r}}}{|\Xi(\boldsymbol{R},t)|^2}, \tag{17.38}$$

and an analogous expression for the$Y$-component. The numerator of Eq. (17.38) (divided by the nuclear mass) is the Cartesian X-component of the nuclear velocity field. The additional term of Eq. (17.36) depending on the gradient of the nuclear phase $S(\boldsymbol{R},t)$ is exactly zero in the chosen gauge. The details of the model Hamiltonian used in this section can be found in Ref. [47].

In the diabatic representation, this model describes the crossing of two similar parabolas, one being slightly displaced both in the X direction and in energy. The adiabatic PESs present a conical intersection at (3.0, 0.0) bohr. The initial nuclear wave packet is taken as Gaussian, with

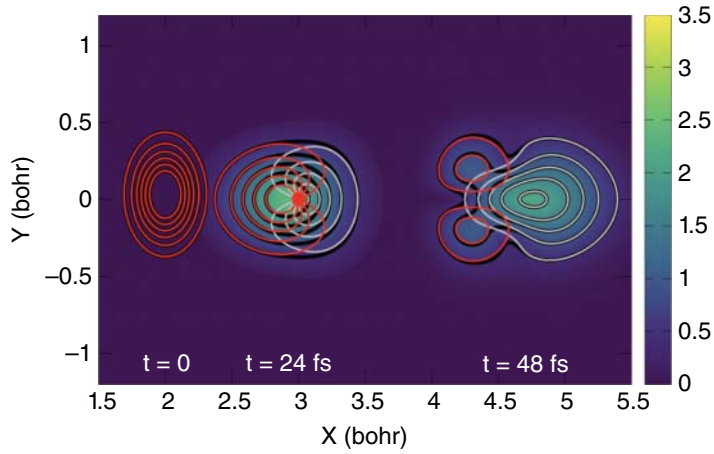


Figure 17.5 Schematic representation of the nuclear wave packet at the beginning of the dynamics ($t = 0$), during the passage through the conical intersection ($t = 24$ fs), and after the non-adiabatic region ($t = 48$ fs). The color-map and the black contour lines show $|\Xi(\boldsymbol{R}, t)|$, whereas the red contour lines represent $|\Xi^{(2)}_{BO}(\boldsymbol{R}, t)|$ and the gray contour lines $|\Xi^{(1)}_{BO}(\boldsymbol{R}, t)|$. A red filled circle indicates the position of the conical intersection. Figure adapted from Ref. [47].

widths $\sigma_X = 0.15$ bohr and $\sigma_Y = 0.197$ bohr, and is centered at $\boldsymbol{R}_0 = (X_0, Y_0) = (2.0, 0.0)$ bohr. The dynamics is initiated in the second adiabatic state (S_1).

Figure 17.5 shows snapshots of the nuclear wave packet in the X, Y-plane. It is initialized on S_1 (red contour lines), slides down towards the conical intersection (the red dot in the figure), transferring amplitudes to S_0 (gray contour lines), and finally separates into two wave packets, one accelerating on the lower surface (gray contour lines) and one moving slower on the more confining upper surface (red contour lines). The dynamics is similar to the case studied in Section 17.3.1, but here it takes place in a two-dimensional space and it involves a funneling process through the conical intersection. The analysis reported here will focus on two main points: first, we are interested in understanding if the TDPES shows peculiar features related to the conical intersection; second, and more importantly, we look for the appearance of any singular behavior of the time-dependent vector potential.

Figure 17.6 shows two views of the TDPES at times $t = 24$ fs (left panel) and $t = 48$ fs (right panel). At $t = 24$ fs we report only a cut of the TDPES along the Y-axis for $X = 3$ bohr, the position of the conical intersection. At this time, the nuclear wave packet is localized around the conical intersection, and in fact we observe two contributions, denoted $|\Xi^{(1)}_{BO}|^2$ and $|\Xi^{(2)}_{BO}|^2$, corresponding to the lower BO PES S_0 and to the upper BO PES S_1, respectively. In this analysis, the GI part of the TDPES has been decomposed into three contributions $\epsilon_{GI} = \epsilon_{GI1} + \epsilon_{GI2} + \epsilon_{GI3}$, that are the three terms in the expression (17.34). The cut of ϵ_{GI1} lies between the adiabatic surfaces, and in this case does not present steps, in fact the two portions of the wave packet are localized in the same region: the TDPES resembles an average potential. Additionally, we observe a peculiar feature of $|\Xi^{(1)}_{BO}|^2$ and $|\Xi^{(2)}_{BO}|^2$, namely peaks at the position of the conical intersection. At this position, there is an infinite coupling between the states S_0 and S_1 (the non-adiabatic coupling diverges here). However, the peaks exactly cancel out, yielding a very smooth [93] (total) nuclear density $|\Xi|^2 = |\Xi^{(1)}_{BO}|^2 + |\Xi^{(2)}_{BO}|^2$. In Figure 17.6, it appears that $\epsilon_{GI2} + \epsilon_{GI3}$ is exactly zero: this is not a general feature of ϵ_{GI2} and ϵ_{GI3}, however, we have observed in various situations that their sum is a smooth function of $\boldsymbol{R}$. Such observation should be kept in mind, since in Section 17.4 those terms will be completely

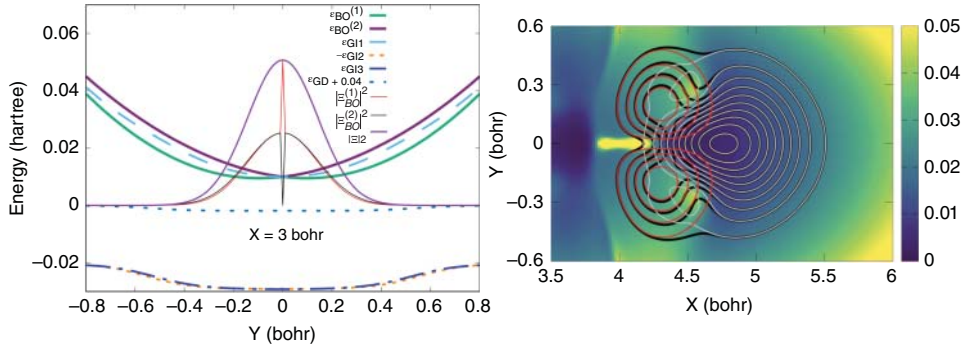


Figure 17.6 Left panel: Representation of all the components forming the TDPES (ϵ_{GI1}, ϵ_{GI2}, ϵ_{GI3}, and ϵ_{GD}) along a cut in the Y direction at $X = 3$ bohr at $t = 24$ fs. The total nuclear probability density ($|\Xi|^2$), adiabatic nuclear probability densities ($|\Xi^{(1)}_{BO}|^2$ and $|\Xi^{(2)}_{BO}|^2$), and adiabatic PESs ($\epsilon^{(1)}_{BO}$ and $\epsilon^{(2)}_{BO}$) are superimposed for comparison. Right panel: Sum of all the GI contributions to the TDPES at $t = 48$ fs (color-map). The black contour lines show $|\Xi(\boldsymbol{R}, t)|$, whereas the red contour lines represent $|\Xi^{(2)}_{BO}(\boldsymbol{R}, t)|$ and the gray contour lines $|\Xi^{(1)}_{BO}(\boldsymbol{R}, t)|$. Figures adapted from Ref. [47].

neglected [55–57, 60] in the approximations introduced to construct the trajectory-based solution of the exact equations (17.15) and (17.16). Finally, in our choice of gauge, the GD part of the TDPES is mostly constant and would lead to a rigid shift of the GI parts of the TDPES.

In Figure 17.6 (right panel) a color-map of the TDPES at $t = 48$ fs is shown[1], along with the contour lines representing the nuclear density (black) and its two adiabatic contributions. The steps observed in a one-dimensional case in Section 17.3.1, leading to a splitting of the nuclear wave packet and explained – in a BO framework – by the S_1 nuclear component separating from the S_0 nuclear component, is clearly visible from the TDPES. The TDPES is composed of two main regions ($3.7 < X < 4.6$ bohr and $4.6 < X < 5.5$ bohr), with a pronounced change of behavior at their interface. A sharp repulsive potential in the X direction composes the first region, supplemented by a central repulsive component at $Y = 0$ bohr. Conversely, the second region only shows a slowly increasing potential towards a larger X value. This strong variation of the TDPES is responsible for the splitting of the nuclear wave packet into two components.

To conclude the analysis of the TDPES, we have observed that the fundamental features of the TDPES highlighted in Section 17.3.1 can also be identified in a two-dimensional case, and can therefore be expected to be retained in more general, higher-dimensional problems. Additionally, the TDPES always appears to be a smooth function of $\boldsymbol{R}$, an appealing property if one is to employ the exact-factorization framework, and thus the TDPES, to numerically simulate non-adiabatic dynamics in electron–nuclear systems.

In the chosen gauge, the time-dependent vector potential equals the nuclear velocity field. Therefore, it encodes information about how the nuclear wave packet moves. The vector potential is shown in Figure 17.7 at times $t = 24$ fs (left panel) and $t = 48$ fs (right panel). The color-code indicates the modulus of the vector field, whereas the arrows indicate the direction. For reference, the position of the nuclear wave packet is represented by the black contour lines.

The time-dependent vector potential exhibits a simple behavior when the wave packet reaches the non-adiabatic region: it mainly points towards larger X, with a strength increasing along X, since the nuclear wave packet itself, after being initiated on S_1, moves rapidly towards the conical-intersection region without spreading significantly in the Y-direction. After the passage

1 We do not include $\epsilon_{GD}(\boldsymbol{R}, t)$ in this sum as it only contributes a nearly constant negative contribution to TDPES.

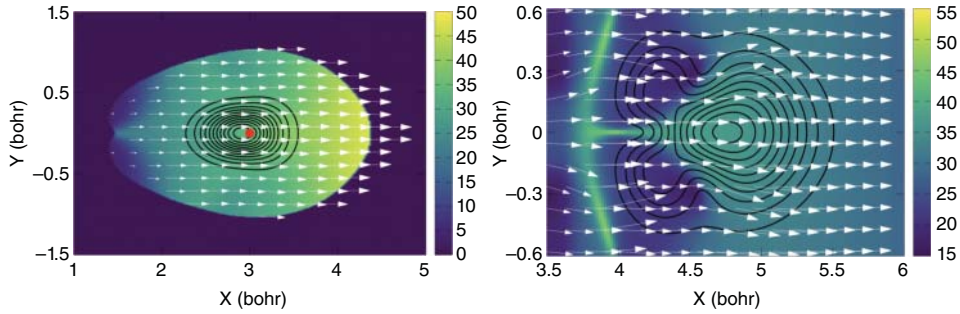


Figure 17.7 Time-dependent vector potential at $t = 24$ fs (left panel) and at $t = 48$ fs (right panel). $\boldsymbol{A}(\boldsymbol{R}, t)$ is represented by arrows, and the color-map represents its magnitude. The black contour lines show $|\Xi(\boldsymbol{R}, t)|$ and the red filled circle the position of the conical intersection. The plot of $\boldsymbol{A}(\boldsymbol{R}, t)$ is restricted to the region of space where $|\Xi(\boldsymbol{R}, t)|^2 > 10^{-10}$. Figures adapted from Ref. [47].

through the conical intersection ($t = 48$ fs), the time-dependent vector potential still mostly points towards larger X, but now with some variations in the region $3.5 < X < 4.5$ bohr. In the region $4.0 < X < 4.5$ bohr, the magnitude of $\boldsymbol{A}(\boldsymbol{R}, t)$ is lower than at larger values of X, reflecting the change in behavior of the nuclear wave packet: the nuclear wave packet component on S_0 overtakes the one on S_1 at $t = 48$ fs, which feels the repulsive potential of S_1. Therefore, the nuclear component evolving on S_1, localized in the region $4.0 < X < 4.5$ bohr, moves slower than the one on S_0, that spreads along larger values of X. In the region $4.5 < X < 6.0$ bohr, the magnitude of $\boldsymbol{A}(\boldsymbol{R}, t)$ starts decreasing, reflecting how the wave packet slows down when it enters a more repulsive part of the PES.

Computing the circulation of the vector potential along a closed path $\mathcal{C}$,

$$\gamma_K(C) = \oint_{\mathcal{C}} \boldsymbol{A}_K(\boldsymbol{R}, t) \cdot d\boldsymbol{R}_K = \int_S \text{curl}\, \boldsymbol{A}_K(\boldsymbol{R}, t) \cdot d\boldsymbol{s} \tag{17.39}$$

we find that, in general, the circulation integral (here denoted $\gamma_K(C)$) is non-zero and depends on the path. The line integral along $\mathcal{C}$ can be replaced by a surface integral according to Stokes' theorem, yielding the flux of the vector field curl $\boldsymbol{A}_K(\boldsymbol{R}, t)$ through S, with S the surface enclosed by $\mathcal{C}$. If $\gamma_K \neq 0$, it means that curl $\boldsymbol{A}_K(\boldsymbol{R}, t) \neq 0$: the time-dependent vector potential $\boldsymbol{A}_K(\boldsymbol{R}, t)$ has non-zero curl, therefore it cannot be written as the gradient of a scalar function (since the curl of a gradient is exactly zero). Recalling the transformation expression of $\boldsymbol{A}_K(\boldsymbol{R}, t)$ under a change of gauge, $\tilde{\boldsymbol{A}}_K(\boldsymbol{R}, t) = \boldsymbol{A}(\boldsymbol{R}, t) + \nabla_K \theta(\boldsymbol{R}, t)$, it is clear that one can set $\tilde{\boldsymbol{A}}_K(\boldsymbol{R}, t) = 0$ if $\boldsymbol{A}_K(\boldsymbol{R}, t) = -\nabla_K \theta(\boldsymbol{R}, t)$. While this is always possible for one-dimensional problems, in higher dimensions the vector potential may have a non-vanishing curl and hence, in general, the vector potential cannot be gauged away.

17.4 Trajectory-Based Solution of the Exact-Factorization Equations

The coupled nuclear (17.15) and electronic (17.16) evolution equations derived from the exact factorization (17.4) of the molecular wave function have been solved based on the representation of the nuclear dynamics in terms of trajectories in Refs. [57, 60, 61, 63]. Note that, as discussed in Ref. [38], the nuclear time-dependent Schrödinger equation (17.15) can be exactly reformulated as two coupled equations, i.e., a Hamilton–Jacobi equation for the phase of the

nuclear wave function Ξ and a continuity equation for the nuclear density $|\Xi|^2$. The former, a partial differential equation that determines the phase of Ξ, can be solved by characteristics (sometimes referred to as *quantum trajectories*). When the coupling between the Hamilton–Jacobi and the continuity equations is neglected (see discussion in Section 17.4.1 below), an approximate solution of Eq. (17.15) is achieved. We denote this procedure "quantum-classical", as the electronic equation, solved at the quantum-mechanical level, is employed to determine the time-dependent potentials, that, in turn, are used to compute the force that generates *classical*, thus approximate, trajectories [38, 55–60, 62, 63]. The idea is to employ the time-dependent potential appearing in the nuclear Schrödinger equation (17.15) to uniquely define the force [58, 60] that generates the trajectories. In the exact factorization the electronic system is not represented in terms of "static" states and multiple potential energy surfaces, but, instead, we have a single time-dependent potential energy surface, thereby overcoming the ambiguity in the definition of the force driving nuclear motion in non-adiabatic conditions. The algorithm has been dubbed coupled-trajectory mixed quantum-classical (CT-MQC). CT-MQC has been employed to simulate non-adiabatic dynamics in model systems [38, 57, 60, 61, 63], and implemented in a developer version of the electronic-structure package CPMD [94], based on (time-dependent) density functional theory (TDDFT), to simulate the photo-induced dynamics in oxirane [34, 56, 62]. We will summarize in Section 17.4.1 the approximations that have led to the derivation of the algorithm, and we will present the results for oxirane in Section 17.4.2.

We will mainly focus on two phenomena occurring in non-adiabatic events, the amplitude transfer between electronic states and quantum decoherence. The former is mainly induced by the non-adiabatic coupling vectors, thus occurring in regions of configuration space where the adiabatic states are close in energy. The latter [55, 95–115] is usually defined as the decay of the off-diagonal elements of the reduced density matrix of the electronic subsystem, a consequence of the coupling with the nuclear environment. Within the BO framework, the electronic subsystem is expected to collapse onto a selected adiabatic state after the dynamics has led the full system through a region of non-adiabatic coupling. It will be argued that decoherence is related to the spatial splitting, or delocalization, of the nuclear wave function.

17.4.1 CT-MQC: The Approximations

Trajectories will be employed to mimic the dynamics of the nuclear wave function $\Xi(\boldsymbol{R}, t)$, that evolves according to the TDSE (17.15). The trajectories can be interpreted as a moving grid, thus we calculate time derivatives "along the flow": partial time derivatives are replaced by total derivatives, using the chain rule $\frac{d}{dt} = \frac{\partial}{\partial t} + \sum_K \boldsymbol{V}_K \cdot \nabla_K$. The velocity of the moving grid point, i.e., of the trajectory, is $\boldsymbol{V}_K = \mathbf{P}_K / M_K$.

Writing the nuclear wave function in polar form, $\Xi(\boldsymbol{R}, t) = |\Xi(\boldsymbol{R}, t)| e^{(i/\hbar) S(\boldsymbol{R}, t)}$, the real part of Eq. (17.15) yields

$$\frac{\partial}{\partial t} S(\boldsymbol{R}, t) = -\sum_{K=1}^{N_n} \frac{[\nabla_K S(\boldsymbol{R}, t) + \boldsymbol{A}_K(\boldsymbol{R}, t)]^2}{2M_K} - \epsilon(\boldsymbol{R}, t) - \sum_{K=1}^{N_n} \frac{-\hbar^2}{2M_K} \frac{\nabla_K^2 |\Xi(\boldsymbol{R}, t)|}{|\Xi(\boldsymbol{R}, t)|}, \tag{17.40}$$

a Hamilton–Jacobi equation in the presence of the time-dependent vector potential $\boldsymbol{A}_K(\boldsymbol{R}, t)$, the TDPES $\epsilon(\boldsymbol{R}, t)$, and of a potential term (the last term in Eq. (17.40)) known in the framework of

2 This derivation is a standard quantum-mechanics exercise. The idea is to insert the polar form of the wave function, i.e., its expression in terms of modulus and phase, into the time-dependent TDSE and to separate the real and imaginary parts. The corresponding Hamilton–Jabobi equation for the phase and continuity equation for the modulus is straightforwardly obtained. The result for the real part is given in Eq. (17.40).

Bohmian dynamics as quantum potential [38]. The imaginary part of Eq. (17.15) yields a continuity equation for the nuclear density. Neglecting the quantum potential, Eq. (17.40) decouples from the continuity equation, and becomes a (standard) classical Hamilton–Jacobi equation. Then, we define the linear momentum as $M_K \mathbf{V}_K = \nabla_K S(\boldsymbol{R}, t) + \boldsymbol{A}_K(\boldsymbol{R}, t) = \mathbf{P}_K$, we introduce the full time derivative of $S(\boldsymbol{R}, t)$, and we apply the spatial derivative $\nabla_{K'}$ on both sides. Equation (17.40) reduces to Newton's equation,

$$\frac{d}{dt}\boldsymbol{P}_K(t)|_{\boldsymbol{R}^{(I)}(t)} = -\nabla_K \left(\epsilon(\boldsymbol{R}, t) + \sum_{K'=1}^{N_n} \boldsymbol{A}_{K'}(\boldsymbol{R}, t) \cdot \frac{\mathbf{P}_{K'}(\boldsymbol{R}, t)}{M_{K'}} \right) + \frac{d}{dt}\boldsymbol{A}_K(\boldsymbol{R}, t)\Big|_{\boldsymbol{R}^{(I)}(t)}, \tag{17.41}$$

i.e., to one of the characteristic (ordinary differential) equations associated to the Hamilton–Jacobi equation (17.40) without the quantum potential (as shown in Ref. [38]). At time t all quantities are evaluated at $\boldsymbol{R}^{(I)}(t)$. The gauge is imposed so as to set to zero the term in parenthesis, $\epsilon(\boldsymbol{R}, t) + \sum_K \boldsymbol{A}_K(\boldsymbol{R}, t) \cdot \mathbf{P}_K(\boldsymbol{R}, t)/M_K = 0$. Equation (17.41) thus becomes $\frac{d}{dt}\boldsymbol{P}_K(t) = \frac{d}{dt}\boldsymbol{A}_K(t)$.

The expression of the TDPES is simplified within CT-MQC by neglecting the terms ϵ_{GI2} and ϵ_{GI3} (remember that in Section 17.3.2 we observed that their sum has a smoother $\boldsymbol{R}$-dependence than ϵ_{GI1}). It is worth noting that both terms have to be neglected in order to maintain gauge invariance. In addition, the approximate expression of the TDPES has to contain the total time derivative of $\Phi_{\boldsymbol{R}}(\boldsymbol{r}, t)$ rather than the partial time derivative.

In order to allow for the interface of CT-MQC with standard quantum-chemistry codes, the BO representation of the electronic wave function is introduced, such that energies, forces and non-adiabatic coupling vectors appear in Eq. (17.16) and can be computed on the fly. The partial time derivative of the electronic wave function in Eq. (17.16) is replaced by the total time derivative, $\frac{\partial}{\partial t}\Phi_{\boldsymbol{R}}(\boldsymbol{r}, t)|_{\boldsymbol{R}^{(I)}(t)} = \frac{d}{dt}\Phi_{\boldsymbol{R}}(\boldsymbol{r}, t) - \sum_{K=1}^{N_n} \frac{\boldsymbol{P}_K(t)}{M_K} \cdot \nabla_K \Phi_{\boldsymbol{R}}(\boldsymbol{r}, t)|_{\boldsymbol{R}^{(I)}(t)}$. These two operations allow us, starting from Eq. (17.16), to derive a set of coupled partial differential equations for the coefficients $C_\alpha(\boldsymbol{R}, t) \to C_\alpha^{(I)}(t)$ – that become functions of the trajectory $\boldsymbol{R}^{(I)}(t)$, and are thus indicated with a superscript (I). In the expression of the electron–nuclear coupling operator $\hat{U}_{en}$ of Eq. (17.19), a dependence on the nuclear wave function appears explicitly. Employing its polar representation,

$$\begin{aligned} \frac{-i\hbar\nabla_K \Xi^{(I)}(t)}{\Xi^{(I)}(t)} + \boldsymbol{A}_\nu^{(I)}(t) &= \left[\nabla_K S^{(I)}(t) + \boldsymbol{A}_K^{(I)}(t)\right] + i\frac{-\hbar\nabla_K |\Xi^{(I)}(t)|}{|\Xi^{(I)}(t)|} \\ &= \boldsymbol{P}_K^{(I)}(t) + i\mathcal{Q}_K^{(I)}(t), \end{aligned} \tag{17.42}$$

we express such dependence in terms of $\boldsymbol{P}_K^{(I)}(t)$, the classical nuclear momentum, and $\mathcal{Q}_K^{(I)}(t)$, the signature quantity of the exact factorization that we have dubbed "quantum momentum".

The only approximation introduced so far in the electronic equation (17.16) has been the neglect of ϵ_{GI2} and ϵ_{GI3} in the TDPES. Additionally, (i) we will neglect all terms that contain products of the quantum momentum and the non-adiabatic coupling vectors, and (ii) we will express the spatial derivatives of the expansion coefficients

$$\nabla_K C_\alpha^{(I)}(t) \simeq \frac{i}{\hbar}\boldsymbol{f}_{\alpha,K}^{(I)}(t) C_\alpha^{(I)}(t) = \frac{i}{\hbar}\left(-\int^t dt' \nabla_K \epsilon_{BO}^{(\alpha),(I)}\right) C_\alpha^{(I)}(t). \tag{17.43}$$

Approximation (i) is justified if the non-adiabatic region is sufficiently localized. Approximation (ii) is discussed in detail in Ref. [57] and is of a more "phenomenological" nature. This approximation is essential to transform the set of partial differential equations for the coefficients in a set of ordinary differential equations, as it allows us to neglect the spatial derivative of the modulus of the coefficients $C_\alpha^{(I)}(t)$ and to express the spatial derivative of the phase in terms of $\boldsymbol{f}_{\alpha,K}^{(I)}(t)$ (time-integral along the trajectory I of the adiabatic forces). Following from Eq. (17.29), the phase of the electronic coefficients $C_\alpha(\boldsymbol{R}, t)$ can be expressed as $S_\alpha(\boldsymbol{R}, t) - S(\boldsymbol{R}, t)$, i.e., the difference between

the phase of the nuclear wave packet $\Xi_{BO}^{(\alpha)}(\boldsymbol{R}, t)$ and the phase of the nuclear wave function $\Xi(\boldsymbol{R}, t)$. Based on semiclassical arguments, the phases are related to the momentum of the moving wave packet, that can be expressed as the time-integral of the force (where the force is the time derivative of the momentum). This approximation has been devised based on previous analysis [41–44, 58] of the expansion coefficients.

In conclusion, the CT-MQC equations can be cast in a very simple form, that is

$$\frac{d}{dt}C_\alpha(t) = \frac{d}{dt}C_\alpha^{\text{Eh.}}(t) + \frac{d}{dt}C_\alpha^{\text{qm}}(t), \tag{17.44}$$

$$\boldsymbol{F}_K = \mathbf{F}_K^{\text{Eh.}} + \mathbf{F}_K^{\text{qm}}, \tag{17.45}$$

where the first terms on the right-hand sides are Ehrenfest-like terms (Eh.), whereas the second terms depend on the quantum momentum (qm). The Ehrenfest-like terms in Eqs. (17.44) and (17.45) are

$$\frac{d}{dt}C_\alpha^{\text{Eh.}}(t) = -\frac{i}{\hbar}\epsilon_{BO}^{(\alpha)}C_\alpha(t) - \sum_{K=1}^{N_n}\boldsymbol{V}_K \cdot \sum_\beta \boldsymbol{d}_{\alpha\beta,K}C_\beta(t) \tag{17.46}$$

$$\boldsymbol{F}_K^{\text{Eh.}} = -\sum_\alpha |C_\alpha(t)|^2 \nabla_K \epsilon_{BO}^{(\alpha)} - \sum_{\alpha,\beta} C_\alpha^*(t)C_\beta(t)\left(\epsilon_{BO}^{(\beta)} - \epsilon_{BO}^{(\alpha)}\right)\mathbf{d}_{\alpha\beta,K}. \tag{17.47}$$

The additional terms, instead, are

$$\frac{d}{dt}C_\alpha^{\text{qm}}(t) = \sum_{K=1}^{N_n}\frac{Q_K}{\hbar M_K}\cdot\left[\boldsymbol{f}_{\alpha,K} - \sum_\beta |C_\beta(t)|^2 \boldsymbol{f}_{\beta,K}\right] C_\alpha(t) \tag{17.48}$$

$$\boldsymbol{F}_K^{\text{qm}} = \sum_\beta |C_\beta(t)|^2 \left(\sum_{K'=1}^{N_n}\frac{2Q_{K'}}{\hbar M_{K'}}\cdot \boldsymbol{f}_{\beta,K'}\right)\left[\boldsymbol{f}_{\beta,K} - \sum_\alpha |C_\alpha(t)|^2 \boldsymbol{f}_{\alpha,K}\right]. \tag{17.49}$$

All $\boldsymbol{R}$-dependent quantities have to be evaluated along a trajectory (I). Equations (17.47) and (17.49) generate the trajectory, that is coupled to n evolution equations ($\alpha, \beta = 1, \ldots, n$) representing the evolution of the electronic system according to Eqs. (17.46) and (17.48).

As is clear from Eq. (17.42), the quantum momentum tracks the spatial variation of the nuclear density, as it contains its spatial derivative. At each time step of the simulated dynamics, the nuclear density has to be reconstructed, for instance by computing a histogram from the distribution of classical trajectories (plus smoothing, by using Gaussian functions localized at the positions of the trajectories [56]). Such calculation requires that at the end of each step of dynamics, the trajectories "communicate" – all at the same time – information about their positions, in order to compute the quantum momentum. Once $Q_K^{(I)}(t)$ is known, the trajectories can perform a new step of dynamics. On the fly calculation of the quantum momentum is possible only if the trajectories are propagated all at the same time, that is why the underlying algorithm has been dubbed "coupled-trajectory"-MQC. It has been shown [55–57, 60–63] that inclusion of the quantum momentum is essential to reproduce quantum decoherence effects.

As an example of the importance of the spatial delocalization of a nuclear wave packet, we show here the nuclear density reconstructed from the distribution of CT-MQC trajectories for a two-state one-dimensional model of extended non-adiabatic coupling with reflection [12, 57, 60, 61, 63]. The adiabatic PESs are plotted in black in Figure 17.8. A Gaussian-shaped wave packet centered at -15 bohr is prepared "on" the lower surface and launched with a positive initial momentum of 20 bohr^{-1} towards the region of non-adiabatic coupling between the two adiabatic states. When the coupling region is overcome, the lower BO PES has a well, whereas the upper BO PES has a barrier. For the chosen initial momentum, the contribution propagating on the lower state is transmitted, while the contribution propagating on the upper state is reflected. In Figure 17.8 the first crossing of the non-adiabatic region of the incoming wave packet (left panel), along with the

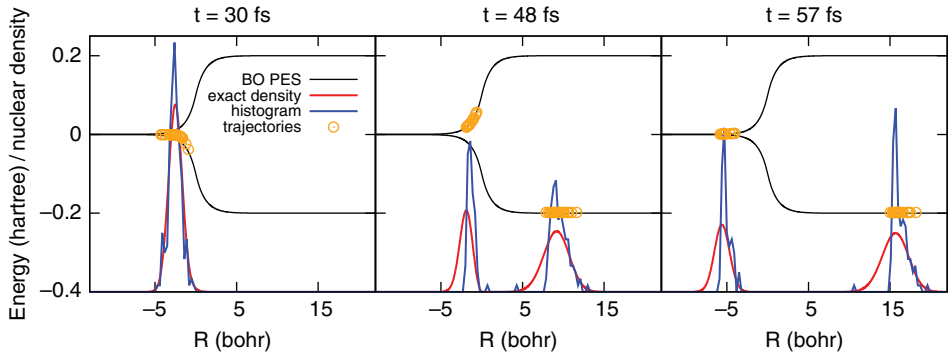


Figure 17.8 Snapshots at $t = 30, 48, 57$ fs of the nuclear density constructed from the distribution of CT-MQC-evolved trajectories (blue lines) for a model of extended non-adiabatic coupling with reflection. CT-MQC results are compared with exact dynamics (red lines). The adiabatic BO PESs are shown in all plots in black, and the positions of the trajectories moving the TDPES (the TDPES is approximated in CT-MQC as described in the main text) are represented as orange dots. The classical distributions are more localized than the quantum-mechanical, an issue that can be corrected by including the quantum potential (last term in Eq. (17.40)) as discussed in Ref. [38]. We remind you here that the quantum potential is not taken into account in CT-MQC, only the quantum momentum is.

splitting of the transmitted and reflected portions of the wave packet (central and right panels) are represented at different time-steps. Here, we compare the histograms constructed from the distribution of classical trajectories (blue lines) with the solution of the full time-dependent Schrödinger equation (red lines). The orange dots show the actual positions of the trajectories that move on the (CT-MQC-approximate form of the) TDPES.

17.4.2 CT-MQC: Photo-Induced Ring Opening in Oxirane

Simulations of the photo-dynamics of oxirane C_2H_4O start at the time in which, upon absorbing a photon, the molecule is excited from its ground electronic state S_0 to the lowest-lying bright state, S_2. The photo-excitation induces rearrangements of the molecular structure that drive the system through a conical intersection between S_2 and S_1 within 7 to 15 fs [56, 62]. This process is illustrated in Figure 17.9, which shows (in the upper panel) the average population of the electronic states, i.e., $\rho_\alpha(t) = N_{traj}^{-1}\sum_I |C_\alpha^{(I)}(t)|^2$, with $N_{traj} = 100$ the number of trajectories ($\alpha = 1, 2, 3$ or, equivalently, S_0, S_1 and S_2). When the molecule reaches the region of the conical intersection, population is transferred from S_2 (fully populated at time $t = 0$) to the first excited state S_1. After that, the dynamics carries on until a conical intersection between S_1 and S_0 is eventually reached (this second event is not discussed here).

Additional information about the S_2-to-S_1 transfer process can be extracted from the analysis of the decoherence indicator $\eta_{23}(t) = N_{traj}^{-1}\sum_I |C_2^{(I)*}(t)C_3^{(I)}(t)|^2$, presented in the lower panel of Figure 17.9. The quantity $\eta_{23}(t)$ (black line in Figure 17.9) shows two pronounced peaks, indication that two groups of trajectories funnel through the S_1/S_2 conical intersection at subsequent times. In order to understand if the time delay between different passages through the conical intersection has an effect on the overall dynamics, the indicator of decoherence has been decomposed in different contributions, associated to different final molecular configurations. We observe four final product structures, (*i*) a right-open ring structure (with probability 36%), shown in the inset representing the molecular structure in the upper left panel of Figure 17.10, (*ii*) a left-open ring structure (with probability 47%), (*iii*) a CC-extended bond structure (with probability 10%), shown in the inset representing the molecular structure in the upper right panel of Figure 17.10, and (*iv*) a closed-ring structure (with probability 7%). All reported values are computed as the

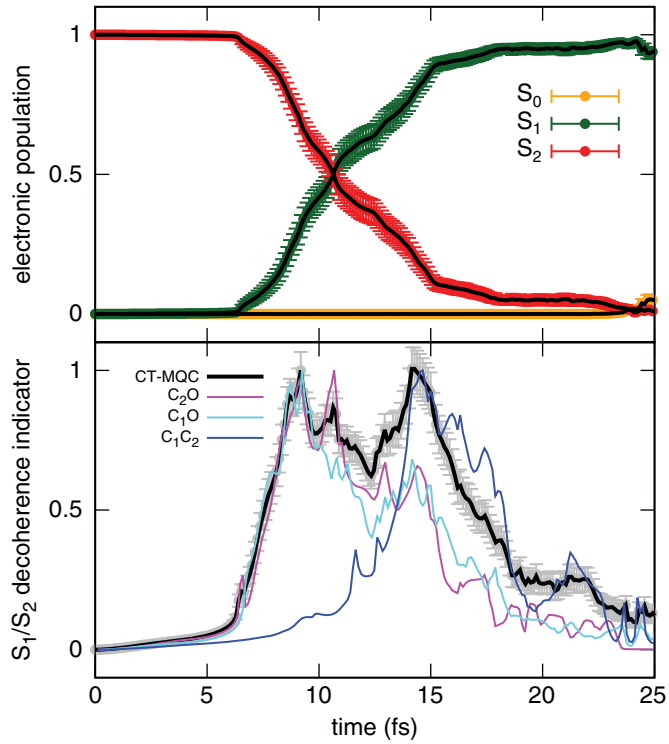


Figure 17.9 Upper panel: electronic populations of S_0 (orange), S_1 (green) and S_2 (red) as functions of time. Lower panel: (normalized) indicator of decoherence for the element S_1/S_2 (black line), and its decomposition in contributions from the three sets of reactive trajectories. The trajectory sets labeled with C_1O (cyan line) and C_2O (magenta line) lead to a final configuration where the oxirane ring opens via the breaking of one of the two equivalent CO bonds; the set of trajectories labeled C_1C_2 (blue line) yields final configurations where the ring opens through the elongation of the CC bond. The error-bars are computed as the standard deviations of the data. Figure adapted from Ref. [62], Copyright (2018) with permission Springer Nature.

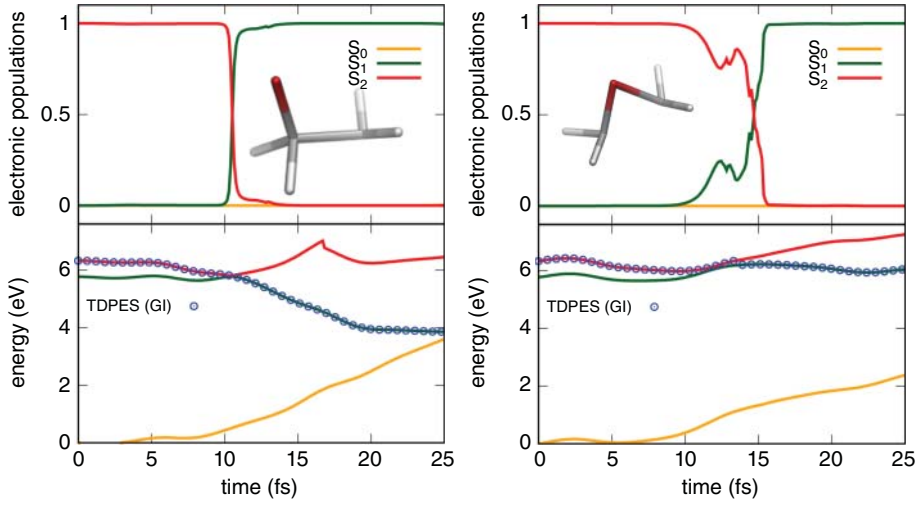


Figure 17.10 Upper panels: populations of the electronic states S_0, S_1, and S_2 as functions of time for two selected trajectories of type (*i*) (left) and of type (*iii*) (right). The color code is the same as used in Figure 17.9. Lower panels: energy profiles (in eV) along the selected trajectories. The zero is set to be the value of the energy of S_0 at time $t = 0$. In the upper panels, oxirane at the final time is shown. Figure adapted from Ref. [62], Copyright (2018) with permission Springer Nature.

ratio between the number of trajectories ending up in the target configuration and the total number of trajectories ($N_{traj} = 100$). Structures (*i*) and (*ii*) are equivalent, and we expect that the observed difference in percentage can be reduced by improving the statistics, thus by increasing the number of simulated trajectories. In these two cases, the ring-opening of oxirane is obtained through the breaking of one of the two CO bonds. In structure (*iii*), the oxirane ring opens via the elongation of the CC bond. A few trajectories, identified as structure (*iv*), are not reactive and remain close to the original molecular configuration. The decoherence indicator is decomposed into contributions associated to the "reactive trajectories". Therefore, we observe that the first peak (cyan and magenta curves in Figure 17.9) between 6 and 12 fs is produced by trajectories that lead to the breakage of one of the two CO bonds. However, these curves do not decay monotonically. Instead, the curves corresponding to the C_1O and C_2O groups both contribute to the second peak (between 12 fs and 17.5 fs), indicating that the first group of trajectories (associated to the first peak) is reached by a second group while funneling through the conical intersection. The main contribution to the second peak (blue line in Figure 17.9) between 12 and 16 fs is produced by trajectories yielding an extended CC bond. These trajectories clearly encounter the non-adiabatic region with some delay when compared to the sets of trajectories (*i*) and (*ii*).

The different reaction channels are clearly a consequence of the topology of the TDPES in configuration space. Therefore, we will now analyze the TDPES for the reactive trajectories of type (*i*) and (*iii*). It is worth recalling that, even though the adiabatic basis has been used to expand the electronic wave function of the exact factorization, the nuclear dynamics is still governed by the TDPES and by the time-dependent vector potential of Eq. (17.15), in their approximate quantum-classical form. The electronic adiabatic basis has been used for practical purposes. Representative trajectories have been selected for the groups (*i*) (Figure 17.10, left) and (*iii*) (Figure 17.10, right). The populations of the electronic states and the adiabatic potential energy for each configuration visited along the trajectories are reported in Figure 17.10 (upper and lower panels, respectively).

The upper panels of Figure 17.10 confirm that the region of strong coupling between states S_2 and S_1 is encountered by trajectories of type (*i*) at earlier times compared to trajectories of type (*iii*). In fact, the populations of the electronic states corresponding to group (*i*) sharply switch at the conical intersection at around 10 fs. This behavior is the consequence of the different shapes of the TDPES, represented as dotted lines in Figure 17.10. The trajectories of group (*iii*) are driven by a TDPES that is initially flat for about 10–15 fs, following the shape of S_2 adiabatic state, until it smoothly approaches and then switches to S_1. Later, these trajectories continue on S_1 without showing a clear tendency to approach the S_1/S_0 conical intersection. In contrast, the TDPES sampled by the trajectories of group (*i*) follow a steeper path that brings to a fast closure the S_2/S_1 gap (within about 15 fs) and subsequently the S_1/S_0 gap (after about 25 fs), suggesting the presence of a funneling process that guides the trajectories to the ground state.

These calculations have been performed with the plane-wave based electronic structure package CPMD [94], employing the PBE [116] functional for ground-state and excited-state calculations. Linear-response TDDFT calculations [117–119] are based on the Tamm–Dancoff approximation [120, 121]. The Kleinman–Bylander [122] pseudo-potential has been used for all atom species together with a plane-wave cutoff of 70 Ry. Initial conditions, i.e., positions and momenta, have been sampled from an ab initio ground-state trajectory of 2 ps at 300 K. The $N_{traj} = 100$ trajectories are propagated with a time step of 0.12 fs (5 a.u.).

17.4.3 CT-MQC: The Algorithm

Currently, for on the fly ab initio calculations, the CT-MQC algorithm is only implemented in a developer version of the CPMD electronic-structure software [94], thus it is not yet released. A code for model-systems calculations based on CT-MQC is available as a module of the E-CAM Software

library [123]. A flowchart describing the actions to be performed in the actual implementation of CT-MQC are given in Figure 17.11. One of the advantages of the CT-MQC algorithm if compared, for instance, with the surface-hopping algorithm [12] is that it is not stochastic. The electronic equation (17.44), for instance, yields the actual populations of the adiabatic states, including decoherence effects. It follows that even a small number of trajectories is able to provide reliable and accurate results. However, as of today, a systematic investigation of the relation between the number of nuclear degrees of freedom of the molecular system of interest and the number of trajectories required for achieving converged results has not yet been conducted.

As for probably most of the trajectory-based approaches to non-adiabatic dynamics, the computational bottleneck still lies in the calculation of electronic structure properties. This bottleneck is even more severe in CT-MQC, since the electronic properties (adiabatic energies and forces, and non-adiabatic coupling vectors) need to be evaluated at each time for all adiabatic states. Furthermore, CT-MQC requires explicit calculations of the non-adiabatic coupling vectors and not only of their products with the classical nuclear momentum, as is the case, for instance, in the surface hopping algorithm. CT-MQC is based on the propagation of coupled trajectories, namely, the N_{traj} trajectories have to be generated simultaneously, which increases the computational cost. The parallelization of the code is, thus, essential.

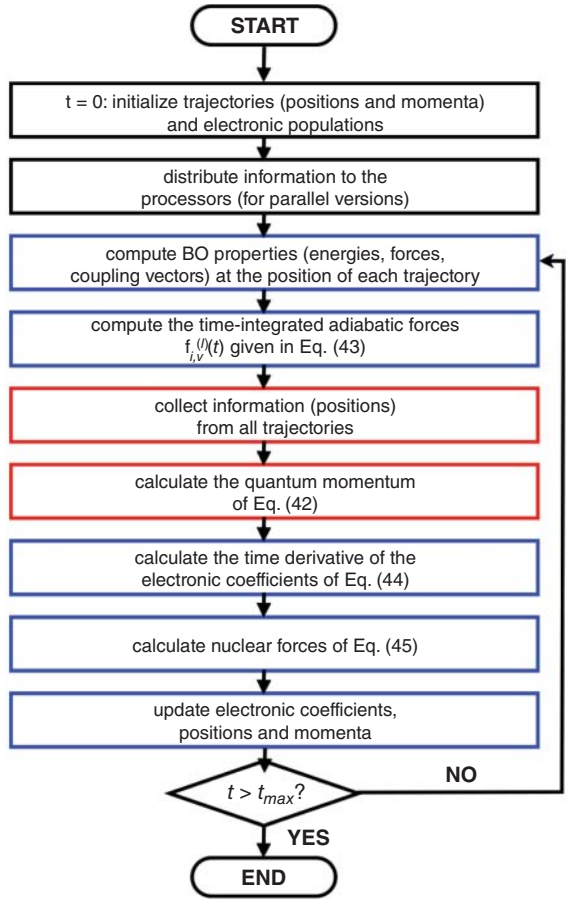


Figure 17.11 Flowchart describing the numerical implementation of the CT-MQC. The blue boxes indicate the operations that each trajectory can perform independently, i.e., the operations that can be parallelized; the red boxes are the operations that cannot be parallelized.

17.5 The Molecular Berry Phase

Before concluding, we briefly present some observations concerning a feature extensively discussed in the literature in relation to adiabatic and non-adiabatic dynamics in molecular systems: the Berry phase. To this end, we focus on the stationary Schrödinger equation for a system of interacting electrons and nuclei, and formulate the time-independent version of the exact factorization.

In order to introduce the concept of the molecular Berry phase, we start from the time-independent version of the adiabatic approximation,

$$\Psi_{\alpha}^{\mathrm{adia}}(\boldsymbol{R},\boldsymbol{r}) = \Xi_{BO}^{(\alpha)}(\boldsymbol{R})\Phi_{\boldsymbol{R}}^{(\alpha)}(\boldsymbol{r}) \tag{17.50}$$

where $\Phi_{\boldsymbol{R}}^{(\alpha)}(\boldsymbol{r})$ is one single BO state. If one variationally optimizes the nuclear wave function $\Xi_{BO}^{(\alpha)}(\boldsymbol{R})$ using the Rayleigh–Ritz principle [124]

$$0 = \frac{\delta}{\delta\Xi_{BO}^{(\alpha)*}(\boldsymbol{R})}\left[\left\langle\Psi_{\alpha}^{\mathrm{adia}}\middle|\hat{H}\middle|\Psi_{\alpha}^{\mathrm{adia}}\right\rangle_{\boldsymbol{R},\boldsymbol{r}} - E_{\alpha}\int d\boldsymbol{R}|\Xi_{BO}^{(\alpha)}(\boldsymbol{R})|^2\right] \tag{17.51}$$

one obtains the following variational equation for $\Xi_{BO}^{(\alpha)}(\boldsymbol{R})$

$$\left[\sum_{K=1}^{N_n}\frac{[-i\hbar\nabla_K + \boldsymbol{A}_{BO,K}^{(\alpha)}(\boldsymbol{R})]^2}{2M_K} + \tilde{\epsilon}_{BO}^{(\alpha)}(\boldsymbol{R})\right]\Xi_{BO}^{(\alpha)}(\boldsymbol{R}) = E_{\alpha}\Xi_{BO}^{(\alpha)}(\boldsymbol{R}), \tag{17.52}$$

with

$$\boldsymbol{A}_{BO,K}^{(\alpha)}(\boldsymbol{R}) = \langle\Phi_{\boldsymbol{R}}^{(\alpha)}|-i\hbar\nabla_K\Phi_{\boldsymbol{R}}^{(\alpha)}\rangle_{\boldsymbol{r}} \tag{17.53}$$

and

$$\tilde{\epsilon}_{BO}^{(\alpha)}(\boldsymbol{R}) = \epsilon_{BO}^{(\alpha)}(\boldsymbol{R}) + \sum_{K=1}^{N_n}\left(\frac{1}{2M_K}\langle\nabla_K\Phi_{\boldsymbol{R}}^{(\alpha)}|\nabla_K\Phi_{\boldsymbol{R}}^{(\alpha)}\rangle_{\boldsymbol{r}} - \frac{[\boldsymbol{A}_{BO,K}^{(\alpha)}(\boldsymbol{R})]^2}{2M_K}\right). \tag{17.54}$$

$\boldsymbol{A}_{BO,K}^{(\alpha)}(\boldsymbol{R})$ is known as *Berry connection* [125]. The loop integral of this object around a conical intersection of the kth BO PES

$$\gamma_{BO,K}^{(\alpha)} = \oint \boldsymbol{A}_{BO,K}^{(\alpha)}(\boldsymbol{R})\cdot d\boldsymbol{R}_K \tag{17.55}$$

may give a non-vanishing Berry phase $\gamma_{BO,K}^{(\alpha)}$. This quantity has been the subject of a huge body of literature [17–20, 22–25, 29, 30, 126–134].

It is worth mentioning here that the time-dependent vector potential of Eq. (17.18) and the TDPES of Eq. (17.20) can be interpreted as time-dependent generalizations of the similar, static, quantities in Eqs. (17.53) and (17.54), appearing within the BO approximation. A formal connection can be established [48, 87], since in the limit of infinite nuclear mass $M \to +\infty$, or equivalently in the limit of the electron–nuclear mass ratio tending to zero $\mu = m_e/M \to 0$, the adiabatic quantities are recovered.

While mathematically perfectly well-defined, there is an intrinsic difficulty when one tries to connect the BO geometric phase with the real world. The difficulty is as follows: Similar to Cauchy's theorem referring to complex-valued functions, the loop integral (17.55) can give a non-vanishing value only if the loop encloses some kind of non-analyticity, such as a conical intersection. If the dependence on $\boldsymbol{R}$ of the function under the integral is smooth, the loop integral is zero. So, on one hand, the sharp feature of conical intersections of BO surfaces is necessary to give a finite $\gamma_{BO,K}^{(k)}$, and, on the other hand, the point of intersection is exactly the point in nuclear configuration space where the BO approximation breaks down because there is no clear separation of electronic and nuclear energy scales. This raises the question: Can we define a quantity analogous to the molecular Berry phase but without making the BO approximation? This is indeed possible by using

the time-independent version [79, 82, 84, 85, 135] of the exact factorization. This allows one to write the full solution of the stationary electron–nuclear Schrödinger equation

$$\hat{H}\Psi(\boldsymbol{R},\boldsymbol{r}) = E\Psi(\boldsymbol{R},\boldsymbol{r}) \tag{17.56}$$

as a single product

$$\Psi(\boldsymbol{R},\boldsymbol{r}) = \Xi(\boldsymbol{R})\Phi_{\boldsymbol{R}}(\boldsymbol{r}), \tag{17.57}$$

with $\int d\boldsymbol{r}|\Phi_{\boldsymbol{R}}(\boldsymbol{r})|^2 = 1$ for each $\boldsymbol{R}$. Note that each αth eigenfunction $\Psi_\alpha(\boldsymbol{R},\boldsymbol{r})$ of the full Hamiltonian, corresponding to the eigenvalue E_α, can be factored in the form (17.57). However, we will not label the electronic and nuclear wave functions with the index α, i.e., we will use the notation $\Psi(\boldsymbol{R},\boldsymbol{r}) = \Xi(\boldsymbol{R})\Phi_{\boldsymbol{R}}(\boldsymbol{r})$ for $\Psi_\alpha(\boldsymbol{R},\boldsymbol{r}) = \Xi^{(\alpha)}(\boldsymbol{R})\Phi^{(\alpha)}_{\boldsymbol{R}}(\boldsymbol{r})$.

Equations similar to those derived in the time-dependent version of the exact factorization are given below, namely the equations defining the nuclear and electronic components of $\Psi(\boldsymbol{r},\boldsymbol{R})$, and the potentials. The nuclear wave function satisfies the stationary Schrödinger equation

$$\left[\sum_{K=1}^{N_n}\frac{[-i\hbar\nabla_K + \boldsymbol{A}_K(\boldsymbol{R})]^2}{2M_K} + \epsilon(\boldsymbol{R})\right]\Xi(\boldsymbol{R}) = E\,\Xi(\boldsymbol{R}), \tag{17.58}$$

while the electronic conditional amplitude is determined by the equation

$$\left[\hat{H}_{BO} + \sum_{K=1}^{N_n}\frac{1}{M_K}\left(\frac{[-i\hbar\nabla_K - \boldsymbol{A}_K]^2}{2} + \left(\frac{-i\hbar\nabla_K\Xi(\boldsymbol{R})}{\Xi(\boldsymbol{R})} + \boldsymbol{A}_K\right)\cdot(-i\hbar\nabla_K - \boldsymbol{A}_K)\right)\right]\Phi_{\boldsymbol{R}}(\boldsymbol{r})$$
$$= \epsilon(\boldsymbol{R})\Phi_{\boldsymbol{R}}(\boldsymbol{r}). \tag{17.59}$$

The nuclear Schrödinger equation (17.58) contains a scalar potential, $\epsilon(\boldsymbol{R})$, which we call *exact PES*. This quantity appears as $\boldsymbol{R}$-dependent eigenvalue in the exact electronic equation (17.59). Furthermore, the nuclear Schrödinger equation contains a vector potential

$$\boldsymbol{A}_K(\boldsymbol{R}) = \langle\Phi_{\boldsymbol{R}}|-i\hbar\nabla_K\Phi_{\boldsymbol{R}}\rangle_{\boldsymbol{r}} \tag{17.60}$$

reminiscent of the Berry connection (17.53) appearing in the adiabatic approximation.

The difference between Eq. (17.53) and Eq. (17.60) is that the electronic wave function appearing in Eq. (17.53) is a BO state while $\Phi_{\boldsymbol{R}}(\boldsymbol{r})$ in Eq. (17.60) is the solution of the exact electronic equation (17.59). Therefore, we call $\boldsymbol{A}_K(\boldsymbol{R})$ the *exact Berry connection* and

$$\gamma_K = \oint \mathbf{A}_K(\boldsymbol{R})\cdot d\boldsymbol{R}_K \tag{17.61}$$

the *exact Berry phase*. It should be noted that γ_K is not a geometric phase in the traditional sense because it does *not* involve the notion of adiabatic parallel transport [19]. Here, adiabatic parallel transport is intended as the slow variation of the parametric dependence of the electronic wave function from $\boldsymbol{R}$ to a neighbor $\boldsymbol{R}'$ with the condition $\mathrm{Im}\,\langle\Phi_{\boldsymbol{R}}|d\Phi_{\boldsymbol{R}'}\rangle_{\boldsymbol{r}} = 0$ [136]. That is the crucial advantage which allows us to define, by Eq. (17.61), a geometric phase without making the adiabatic approximation. Hence, the latter is potentially measurable in experiments [86]. How the BO Berry phase (17.55) compares to the exact concept (17.61) has been investigated in two different model systems for which the full electron–nuclear problem can be exactly numerically solved. Both model systems refer to triangular molecules. The first one [87] is the Shin–Metiu model of Section 17.3.1 extended to two dimensions. The relevant nuclear degrees of freedom are vibrational modes. In the second model system [88], the total energy of the complete electron–nuclear system is highly degenerate and the relevant nuclear mode is a pseudo-rotation.

In the BO approximation, the Shin–Metiu model yields a standard Berry phase of value π. In the second model, one also obtains a Berry phase in the BO approximation, but its value lies in the interval $[0,\pi]$ and the specific value depends on the state, i.e., on the choice of linear combination

in the degenerate subspace. In both cases, the BO Berry phase does not depend on the path chosen for the loop integral in Eq. (17.55) as long as it encloses the conical intersection. If one calculates the exact phase in Eq. (17.61), the exact solution of the Shin–Metiu model yields the value zero [87]. Hence, for this particular case, one can rightfully say that the BO Berry phase is an artifact. In view of this result, one might be tempted to believe that the exact phase (17.61) always vanishes. This, however, is not the case. In the model where the relevant nuclear mode is a pseudo-rotation, the exact phase (17.61) can be non-zero. Its value also lies in the interval $[0, \pi]$ (as in the BO approximation) but the phase is not a topological phase, i.e., its value depends on the path chosen in the loop integral (17.61). A detailed analysis is given in Ref. [88].

An interesting aspect arises from the fact that the BO approximation can be viewed as infinite-nuclear-mass limit of the full problem [48]. How can it be that for any finite value of the nuclear mass the Shin–Metiu model yields zero for the exact Berry phase (17.61) while the infinite nuclear mass limit, i.e., the BO approximation, yields the value of π? Knowing that a non-zero Berry phase can only come from some non-analyticity in the $\boldsymbol{R}$-dependence of the electronic wave function $\Phi_{\boldsymbol{R}}(\boldsymbol{r})$, it is useful to study this $\boldsymbol{R}$-dependence as a function of nuclear mass. For the purpose of visualization only, we do not look at the wave function $\Phi_{\boldsymbol{R}}(\boldsymbol{r})$ directly (simply because it depends on four coordinates and is, therefore, difficult to plot) but instead we visualize a derived quantity, namely

$$\boldsymbol{D}(\boldsymbol{R}) = \int d\boldsymbol{r}\, \boldsymbol{r}\, \Phi_{\boldsymbol{R}}(\boldsymbol{r}). \tag{17.62}$$

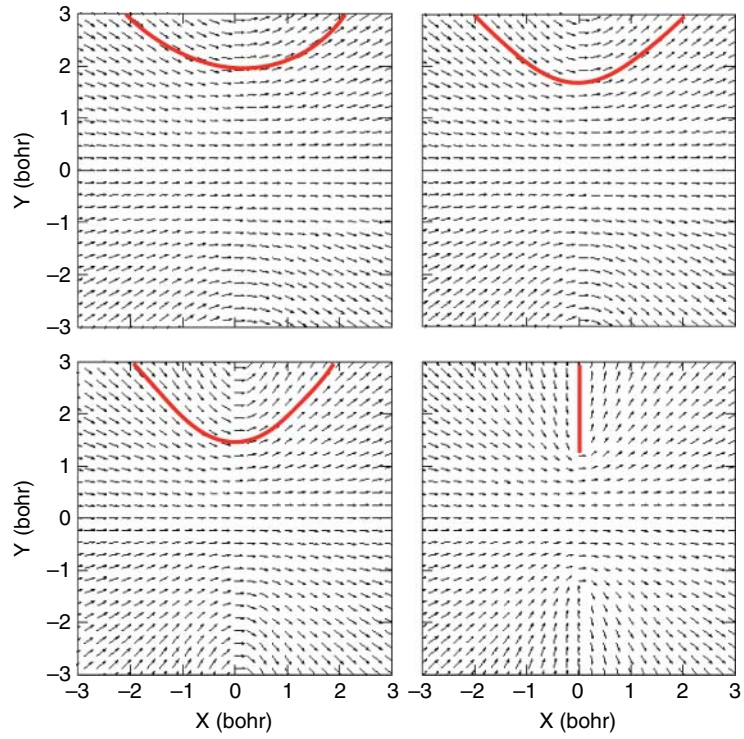


Figure 17.12 Plots of the vector field $\boldsymbol{D}(\boldsymbol{R}) = \boldsymbol{D}(X, Y)$ for different values of the nuclear mass $M = 10\ m_e, 20\ m_e, 50\ m_e, +\infty$ in the two-dimensional Shin–Metiu model of Ref. [60]. In the BO limit, for $M = +\infty$, the electronic wave function used in Eq. (17.62) is the first excited state of $\hat{H}_{BO}$. The red lines are a guide for the eyes.

Note that $\Phi_{\boldsymbol{R}}(\boldsymbol{r})$ is the solution of Eq. (17.59). For each nuclear configuration $\boldsymbol{R}$, it has a change of sign as function of $\boldsymbol{r}$ in our particular model system. Therefore, $\boldsymbol{D}(\boldsymbol{R})$ visualizes the wave function, at each $\boldsymbol{R}$, as an arrow that points from the positive to the negative values of $\Phi_{\boldsymbol{R}}(\boldsymbol{r})$. Figure 17.12 shows the two-dimensional vector field $\boldsymbol{D}(\boldsymbol{R}) = \boldsymbol{D}(X, Y)$ for four different values of the nuclear mass, increasing from the upper left to upper right, to lower left to lower right panels. The lower right panel corresponds to the infinite nuclear mass limit. While for any finite value of the nuclear mass the vector field $\boldsymbol{D}(\boldsymbol{R})$ is perfectly smooth, and hence leads to a vanishing Berry phase, the $M \to +\infty$ limit shows a discontinuous jump at *12 o'clock*. This non-analyticity explains the value π for the BO Berry phase. The plot clearly shows that the BO limit is a singular limit. For any real-world nuclear mass $M < +\infty$, there is no Berry phase for this model system.

17.6 Conclusions

In this chapter we have presented the exact factorization of the electron–nuclear wave function. The basic theory has been discussed, along with the connection between the exact-factorization perspective on non-adiabatic dynamics and the, more standard, Born–Oppenheimer representation. In doing so, we have introduced the concepts of time-dependent vector potential and time-dependent potential energy surface, that are the potentials accounting for electronic excited-state effects on nuclear dynamics. The analysis of the time-dependent potentials has been based on one- and two-dimensional model systems, for which exact quantum-mechanical results are easily accessible. The features of the time-dependent potentials have been presented in typical non-adiabatic situations, namely for the photo-activated dynamics of nuclear wave packets that relax through an avoided crossing or through a conical intersection. We have also shown how those potentials can be determined within a numerical scheme that describes nuclear dynamics in terms of classical trajectories, and that solves electronic dynamics via determining adiabatic energies, forces and couplings on the fly based on standard quantum-chemistry calculations. The coupled-trajectory mixed quantum-classical, CT-MQC, algorithm derived from the exact factorization has been applied to simulate the photo-induced ring-opening processes in oxirane. Finally, we have briefly discussed how to translate the concept of the molecular Berry phase, arising in the context the BO treatment of the stationary electron–nuclear problem, to the language of the exact factorization. Interesting differences between the two approaches have emerged, which are the subject of ongoing investigations.

Acknowledgments

E.K.U.G. acknowledges financial support by the European Research Council Advanced Grant FACT (ERC-2017-AdG-788890)

References

1 Worth, G. and Lasorne, B. (2020). *Gaussian Wave packets and the DD-vMCG Approach*. Wiley.
2 Curchod, B.F.E. (2020). *Full and Ab Initio Multiple Spawning*. Wiley.
3 Kirrander, A. and Vacher, M. (2020). *Ehrenfest Methods for Electron and Nuclear Dynamics*. Wiley.
4 Mai, S., Marquetand, P., and Gonzalez, L. (2020). *Surface-Hopping Molecular Dynamics*. Wiley.

5 Albareda, G. and Tavernelli, I. (2020). *Bohmian Approaches to Nonadiabatic Molecular Dynamics*. Wiley.

6 Conte, R. and Ceotto, M. (2020). *Semiclassical Molecular Dynamics for Spectroscopic Calculations*. Wiley.

7 Saller, M.A.C., Runeson, J.E., and Richardson, J.O. (2020). *Path Integral Approaches to Nonadiabatic Dynamics*. Wiley.

8 Born, M. and Oppenheimer, R.J. (1927). Zur Quantentheorie der Molekeln. *Ann. Phys.* 389: 457–484.

9 Cafiero, M., Bubin, S., and Adamowicz, L. (2003). Non-Born-Oppenheimer calculations of atoms and molecules. *Phys. Chem. Chem. Phys.* 5: 1491–1501.

10 Bubin, S., Pavanello, M., Tung, W.C. et al. (2013). Born-Oppenheimer and Non-Born-Oppenheimer, atomic and molecular calculations with explicitly correlated gaussians. *Chem. Rev.* 113: 36–79.

11 Mátyus, E. (2018). Pre-Born-Oppenheimer molecular structure theory. *Mol. Phys.* 117: 590–609.

12 Tully, J.C. (1990). Molecular dynamics with electronic transitions. *J. Chem. Phys.* 93: 1061–1071.

13 Tully, J.C. (1998). Mixed quantum classical dynamics. *Faraday Discuss.* 110: 407–419.

14 Ben-Nun, M. and Martínez, T.J. (1998). Nonadiabatic molecular dynamics: validation of the multiple spawning method for a multidimensional problem. *J. Chem. Phys.* 108: 7244–7257.

15 Michl, J. and Bonačić-Koutecký, V. (1990). *Electronic Aspects of Organic Photochemistry*. New York: Wiley.

16 Michl, J. (1972). Photochemical reactions of large molecules. I. A simple physical model of photochemical reactivity. *Mol. Photochem.* 4: 243–257.

17 Mead, C.A. and Truhlar, D.G. (1979). On the determination of Born–Oppenheimer nuclear motion wave functions including complications due to conical intersections and identical nuclei. *J. Chem. Phys.* 70 (5): 2284–2296.

18 Berry, M.V. and Wilkinson, M. (1984). Diabolical points in the spectra of triangles. *Proc. Roy. Soc. London Ser. A.* 392 (1802): 15–43.

19 Berry, M.V. (1984). Quantal phase factors accompanying adiabatic changes. *Proc. Roy. Soc. London Ser. A.* 392 (1802): 45–57.

20 Yarkony, D.R. (1996). Diabolical conical intersections. *Rev. Mod. Phys.* 68 (4): 985–1013.

21 Bernardi, F., Olivucci, M., and Robb, M.A. (1996). Potential energy surface crossings in organic photochemistry. *Chem. Soc. Rev.* 25 (5): 321–328.

22 Yarkony, D.R. (2001). Conical intersections: the new conventional wisdom. *J. Phys. Chem.* 105: 6277–6293.

23 Domcke, W., Yarkony, D., and Köppel, H. (eds.) (2004). *Conical Intersections: Electronic Structure, Dynamics & Spectroscopy 15*. World Scientific Pub Co Inc.

24 Worth, G.A. and Cederbaum, L.S. (2004). Beyond Born-Oppenheimer: molecular dynamics through a conical intersection. *Annu. Rev. Phys. Chem.* 55: 127–158.

25 Baer, M. (2006). *Beyond Born-Oppenheimer: Electronic Nonadiabatic Coupling Terms and Conical Intersections*. Wiley.

26 Levine, B.G. and Martínez, T.J. (2007). Isomerization through conical intersections. *Annu. Rev. Phys. Chem.* 58: 613–634.

27 Martinez, T.J. (2010). Physical chemistry: seaming is believing. *Nature* 467 (7314): 412–413.

28 Matsika, S. and Krause, P. (2011). Nonadiabatic events and conical intersections. *Annu. Rev. Phys. Chem.* 62: 621–643.

29 Domcke, W. and Yarkony, D.R. (2012). Role of conical intersections in molecular spectroscopy and photoinduced chemical dynamics. *Annu. Rev. Phys. Chem.* 63: 325–352.

30 Domcke, W., Yarkony, D., and Köppel, H. (2012). *Conical Intersections: Theory, Computation and Experiment 17*. World Scientific Pub Co Inc.

31 Malhado, J.P., Bearpark, M.J., and Hynes, J.T. (2014). Non-adiabatic dynamics close to conical intersections and the surface hopping perspective. *Front. Chem.* 2: 97.

32 Zhu, X. and Yarkony, D.R. (2016). Non-adiabaticity: the importance of conical intersections. *Mol. Phys.* 114 (13): 1983–2013.

33 Xie, C., Malbon, C.L., Yarkony, D.R. et al. (2018). Signatures of a conical intersection in adiabatic dissociation on the ground electronic state. *J. Am. Chem. Soc.*: 1986–1989.

34 Agostini, F., Curchod, B.F.E., Vuilleumier, R. et al. (2018). TDDFT and quantum-classical dynamics: a universal tool describing the dynamics of matter. In: *Handbook of Materials Modeling* (eds. W. Andreoni and S. Yip), 1–47. Netherlands: Springer.

35 Abedi, A., Maitra, N.T., and Gross, E.K.U. (2010). Exact factorization of the time-dependent electron–nuclear wave function. *Phys. Rev. Lett.* 105 (12): 123002.

36 Abedi, A., Maitra, N.T., and Gross, E.K.U. (2012). Correlated electron–nuclear dynamics: exact factorization of the molecular wave-function. *J. Chem. Phys.* 137 (22): 22A530.

37 Agostini, F. and Curchod, B.F.E. (2019). Different flavors of non-adiabatic molecular dynamics. *WIREs Comput. Mol. Sci.* 9: e1417.

38 Agostini, F., Tavernelli, I., and Ciccotti, G. (2018). Nuclear quantum effects in electronic (non)adiabatic dynamics. *Eur. Phys. J. B.* 91: 139.

39 Alonso, J.L., Clemente-Gallardo, J., Echeniche-Robba, P., and Jover-Galtier, J.A. (2013). Comment on Correlated electron–nuclear dynamics: exact factorization of the molecular wave-function. *J. Chem. Phys.* 139: 087101.

40 Abedi, A., Maitra, N.T., and Gross, E.K.U. (2013). Reply to comment on Correlated electron–nuclear dynamics: exact factorization of the molecular wave-function. *J. Chem. Phys.* 139 (8): 087102.

41 Abedi, A., Agostini, F., Suzuki, Y., and Gross, E.K.U. (2013). Dynamical steps that bridge piecewise adiabatic shapes in the exact time-dependent potential energy surface. *Phys. Rev. Lett.* 110 (26): 263001.

42 Agostini, F., Abedi, A., Suzuki, Y., and Gross, E.K.U. (2013). Mixed quantum-classical dynamics on the exact time-dependent potential energy surfaces: a novel perspective on non-adiabatic processes. *Mol. Phys.* 111 (22-23): 3625–3640.

43 Agostini, F., Min, S.K., and Gross, E.K.U. (2015). Semiclassical analysis of the electron–nuclear coupling in electronic non-adiabatic processes. *Ann. Phys.* 527 (9-10): 546–555.

44 Agostini, F., Abedi, A., Suzuki, Y. et al. (2015). The exact forces on classical nuclei in non-adiabatic charge transfer. *J. Chem. Phys.* 142 (8): 084303.

45 Curchod, B.F.E., Agostini, F., and Gross, E.K.U. (2016). An exact factorization perspective on quantum interferences in non-adiabatic dynamics. *J. Chem. Phys.* 145: 034103.

46 Curchod, B.F.E. and Agostini, F. (2017). On the dynamics through a conical intersection. *J. Phys. Chem. Lett.* 8: 831–837.

47 Agostini, F. and Curchod, B.F.E. (2018). When the exact factorization meets conical intersections. *Eur. Phys. J. B.* 91: 141.

48 Eich, F.G. and Agostini, F. (2016). The adiabatic limit of the exact factorization of the electron–nuclear wave function. *J. Chem. Phys.* 145: 054110.

49 Scherrer, A., Agostini, F., Sebastiani, D. et al. (2015). Nuclear velocity perturbation theory for vibrational circular dichroism: an approach based on the exact factorization of the electron–nuclear wave function. *J. Chem. Phys.* 143 (7): 074106.

50 Scherrer, A., Agostini, F., Sebastiani, D. et al. (2017). On the mass of atoms in molecules: beyond the Born-Oppenheimer approximation. *Phys. Rev. X* 7: 031035.

51 Schild, A., Agostini, F., and Gross, E.K.U. (2016). Electronic flux density beyond the born-oppenheimer approximation. *J. Phys. Chem. A* 120: 3316.

52 deCarvalho, F.F., Bouduban, M.E.F., Curchod, B.F.E., and Tavernelli, I. (2014). Nonadiabatic molecular dynamics based on trajectories. *Entropy* 16 (1): 62–85.

53 Suzuki, Y. and Watanabe, K. (2016). Bohmian mechanics in the exact factorization of electron–nuclear wave functions. *Phys. Rev. A* 94: 032517.

54 Suzuki, Y., Abedi, A., Maitra, N.T., and Gross, E.K.U. (2015). Laser-induced electron localization in H_2^+: mixed quantum-classical dynamics based on the exact time-dependent potential energy surface. *Phys. Chem. Chem. Phys.* 17: 29271–29280.

55 Ha, J.-K., Lee, I.S., and Min, S.K. (2018). Surface hopping dynamics beyond non-adiabatic couplings for quantum coherence. *J. Phys. Chem. Lett.* 9: 1097–1104.

56 Min, S.K., Agostini, F., Tavernelli, I., and Gross, E.K.U. (2017). Ab initio non-adiabatic dynamics with coupled trajectories: a rigorous approach to quantum (de)coherence. *J. Phys. Chem. Lett.* 8: 3048–3055.

57 Agostini, F., Min, S.K., Abedi, A., and Gross, E.K.U. (2016). Quantum-classical non-adiabatic dynamics: coupled- vs. independent-trajectory methods. *J. Chem. Theory Comput.* 12 (5): 2127–2143.

58 Agostini, F., Abedi, A., and Gross, E.K.U. (2014). Classical nuclear motion coupled to electronic non-adiabatic transitions. *J. Chem. Phys.* 141 (21): 214101.

59 Abedi, A., Agostini, F., and Gross, E.K.U. (2014). Mixed quantum-classical dynamics from the exact decomposition of electron–nuclear motion. *Europhys. Lett.* 106 (3): 33001.

60 Min, S.K., Agostini, F., and Gross, E.K.U. (2015). Coupled-trajectory quantum-classical approach to electronic decoherence in non-adiabatic processes. *Phys. Rev. Lett.* 115 (7): 073001.

61 Agostini, F. (2018). An exact-factorization perspective on quantum-classical approaches to excited-state dynamics. *Eur. Phys. J. B.* 91: 143.

62 Curchod, B.F.E., Agostini, F., and Tavernelli, I. (2018). CT-MQC – A coupled-trajectory mixed quantum/classical method including non-adiabatic quantum coherence effects. *Eur. Phys. J. B.* 91: 168.

63 Gossel, G., Agostini, F., and Maitra, N.T. (2018). Coupled-trajectory mixed quantum-classical algorithm: a deconstruction. *J. Chem. Theory Comput.* 14: 4513–4529.

64 Requist, R. and Gross, E.K.U. (2016). Exact factorization-based density functional theory of electrons and nuclei. *Phys. Rev. Lett.* 117: 193001.

65 Li, C., Requist, R., and Gross, E.K.U. (2018). Density functional theory of electron transfer beyond the Born–Oppenheimer approximation: case study of LiF. *J. Chem. Phys.* 148: 084110.

66 Requist, R., Proetto, C.R., and Gross, E.K.U. (2019). Exact factorization-based density functional theory of electron-phonon systems. *Phys. Rev. B* 99: 165136.

67 Khosravi, E., Abedi, A., Rubio, A., and Maitra, N.T. (2017). Electronic non-adiabatic dynamics in enhanced ionization of isotopologues of hydrogen molecular ions from the exact factorization perspective. *Phys. Chem. Chem. Phys.* 19: 8269–8281.

68 Khosravi, E., Abedi, A., and Maitra, N.T. (2015). Exact potential driving the electron dynamics in enhanced ionization of H_2^+. *Phys. Rev. Lett.* 115: 263002.

69 Suzuki, Y., Abedi, A., Maitra, N.T. et al. (2014). Electronic Schrödinger equation with nonclassical nuclei. *Phys. Rev. A* 89 (4): 040501(R).

70 Schild, A. and Gross, E.K.U. (2017). Exact single-electron approach to the dynamics of molecules in strong laser fields. *Phys. Rev. Lett.* 118: 163202.

71 Hoffmann, N.M., Appel, H., Rubio, A., and Maitra, N. (2018). Light-matter interactions via the exact factorization approach. *Eur. Phys. J. B.* 91: 180.

72 Abedi, A., Khosravi, E., and Tokatly, I. (2018). Shedding light on correlated electronphoton states using the exact factorization. *Eur. Phys. J. B.* 91: 194.

73 Jeck, T., Sutcliffe, B.T., and Woolley, R.G. (2015). On factorization of molecular wave functions. *J. Phys. A Math. Theor.* 48: 445201.

74 Parashar, S., Sajeev, Y., and Ghosh, S.K. (2015). Towards an exact factorization of the molecular wave function. *Mol. Phys.* 113 (19-20): 3067–3072.

75 Bishop, D.M. and Cheung, L.M. (1977). Conditional probability amplitudes and the variational principle. *Chem. Phys. Lett.* 50 (1): 172–174.

76 Hunter, G. and Tai, C.C. (1982). Variational marginal amplitudes. *Int. J. Quantum Chem.* 21: 1041–1050.

77 Hunter, G. (1981). Nodeless wave functions and spiky potentials. *Int. J. Quantum Chem.* 19: 755–761.

78 Hunter, G. (1975). Conditional probability amplitudes in wave mechanics. *Int. J. Quantum Chem.* 9: 237–242.

79 Lefebvre, R. (2015). Perturbations in vibrational diatomic spectra: factorization of the molecular wave function. *J. Chem. Phys.* 142 (7): 074106.

80 Lefebvre, R. (2015). Factorized molecular wave functions: analysis of the nuclear factor. *J. Chem. Phys.* 142 (21): 214105.

81 Cederbaum, L.S. (2015). The exact wave function of interacting N degrees of freedom as a product of N single-degree-of-freedom wave functions. *Chem. Phys.* 457 (22): 129.

82 Cederbaum, L.S. (2013). The exact molecular wave function as a product of an electronic and a nuclear wave function. *J. Chem. Phys.* 138 (22): 224110.

83 Chiang, Y.-C., Klaiman, S., Otto, F., and Cederbaum, L.S. (2014). The exact wave function factorization of a vibronic coupling system. *J. Chem. Phys.* 140 (5): 054104.

84 Gidopoulos, N.I. and Gross, E.K.U. (2014). Electronic non-adiabatic states: towards a density functional theory beyond the Born-Oppenheimer approximation. *Phil. Trans. R. Soc. A* 372: 20130059.

85 Gidopoulos, N.I. and Gross, E.K.U (2005). Electronic non-adiabatic states. *arXiv:cond-mat/0502433*.

86 Requist, R., Proetto, C.R., and Gross, E.K.U. (2017). Asymptotic analysis of the Berry curvature in the E $\otimes$ e Jahn-Teller model. *Phys. Rev. A* 96: 062503.

87 Min, S.K., Abedi, A., Kim, K.S., and Gross, E.K.U. (2014). Is the molecular berry phase an artefact of the Born-Oppenheimer approximation? *Phys. Rev. Lett.* 113 (26): 263004.

88 Requist, R., Tandetzky, F., and Gross, E.K.U. (2016). Molecular geometric phase from the exact electron–nuclear factorization. *Phys. Rev. A* 93: 042108.

89 Shin, S. and Metiu, H. (1995). Nonadiabatic effects on the charge transfer rate constant: a numerical study of a simple model system. *J. Chem. Phys.* 102: 9285–9295.

90 Feit, M.D., Fleck, F.A., and Steiger, A. (1982). Solution of the Schrödinger equation by a spectral method. *J. Comput. Phys.* 47 (3): 412–433.

91 Gossel, G.H., Lacombe, L., and Maitra, N.T. (2019). On the numerical solution of the exact factorization equations. *J. Chem. Phys.* 150: 154112.

92 Agostini, F., Gross, E.K.U., and Curchod, B.F.E. (2019). Electron-nuclear entanglement in the time-dependent molecular wave function. *Comput. Theo. Chem.*: 99, 1151–106.

93 Hader, K., Albert, J., Gross, E.K.U., and Engel, V. (2017). Electron-nuclear wave-packet dynamics through a conical intersection. *J. Chem. Phys.* 146: 074304.

94 CPMD, http://www.cpmd.org/, Copyright IBM Corp 1990-2015, Copyright MPI für Festkörperforschung Stuttgart 1997–2001.

95 Engel, G.S., Calhoun, T.R., Read, E.L. et al. (2007). Evidence for wavelike energy transfer through quantum coherence in photosynthetic systems. *Nature* 446: 782–786.

96 Ishizaki, A. and Fleming, G.R. (2009). Theoretical examination of quantum coherence in a photosynthetic system at physiological temperature. *Proc. Natl. Acad. Sci. USA* 106: 17255–17260.

97 Scholes, G.D. (2010). Quantum coherent electronic energy transfer: did nature think of it first? *J. Phys. Chem. Lett.* 1: 2–8.

98 Collini, E., Wong, C.Y., Wilk, K.E. et al. (2010). Coherently wired light-harvesting in photosynthetic marine algae at ambient temperature. *Nature* 463: 644–647.

99 Panitchayangkoon, G., Hayes, D., Fransted, K.A. et al. (2010). Long-lived quantum coherence in photosynthetic complexes at physiological temperature. *Proc. Natl. Acad. Sci. U.S.A.* 107: 12766–12770.

100 Turner, D.B., Wilk, K.E., Curmi, P.M.G., and Scholes, G.D. (2011). Comparison of electronic and vibrational coherence measured by two-dimensional electronic spectroscopy. *J. Phys. Chem. Lett.* 2: 1904–1911.

101 Huo, P. and Coker, D.F. (2011). Theoretical study of coherent excitation energy transfer in cryptophyte phycocyanin 645 at physiological temperature. *J. Phys. Chem. Lett.* 2: 825–833.

102 Strümpfer, J., Sener, M., and Schulten, K. (2012). How quantum coherence assists photosynthetic light-harvesting. *J. Phys. Chem. Lett.* 3: 536–542.

103 Dorfman, K.E., Voronine, D.V., Mukamel, S., and Scully, M.O. (2013). Photosynthetic reaction center as a quantum heat engine. *Proc. Natl. Acad. Sci. U.S.A.* 110: 2746–2751.

104 Rozzi, C.A., Falke, S.M., Spallanzani, N. et al. (2013). Quantum coherence controls the charge separation in a prototypical artificial light-harvesting system. *Nature Comm.* 4: 1602.

105 Akimov, A.V. and Prezhdo, O.V. (2013). Persistent electronic coherence despite rapid loss of electron–nuclear correlation. *J. Phys. Chem. Lett.* 4: 3857–3864.

106 Landry, B.R. and Subotnik, J.E. (2014). Quantifying the lifetime of triplet energy transfer processes in organic chromophores: a case study of 4-2-(Naphthylmethyl)benzaldehyde. *J. Chem. Theory Comput*. 10: 4253–4263.

107 Curutchet, C. and Mennucci, B. (2017). Quantum chemical studies of light harvesting. *Chem. Rev.* 117: 294–343.

108 Schwartz, B.J., Bittner, E.R., Prezhdo, O.V., and Rossky, P.J. (1996). Quantum decoherence and the isotope effect in condensed phase non-adiabatic molecular dynamics simulations. *J. Chem. Phys.* 104: 5942–5955.

109 Fang, J.-T. and Hammes-Schiffer, S. (1999). Improvement of the internal consistency in trajectory surface hopping. *J. Phys. Chem. A* 103: 9399–9407.

110 Granucci, G. and Persico, M. (2007). Critical appraisal of the fewest switches algorithm for surface hopping. *J. Chem. Phys.* 126: 134114.

111 Shenvi, N., Subotnik, J.E., and Yang, W. (2011). Simultaneous-trajectory surface hopping: a parameter-free algorithm for implementing decoherence in non-adiabatic dynamics. *J. Chem. Phys.* 134: 144102.

112 Shenvi, N., Subotnik, J.E., and Yang, W. (2011). Phase-corrected surface hopping: correcting the phase evolution of the electronic wave function. *J. Chem. Phys.* 13: 024101.

113 Shenvi, N. and Yang, W. (2012). Achieving partial decoherence in surface hopping through phase correction. *J. Chem. Phys.* 137: 22A528.

114 Subotnik, J.E. and Shenvi, N. (2011). A new approach to decoherence and momentum rescaling in the surface hopping algorithm. *J. Chem. Phys.* 134: 024105.

115 Subotnik, J.E. and Shenvi, N. (2011). Decoherence and surface hopping: when can averaging over initial conditions help capture the effects of wave packet separation? *J. Chem. Phys.* 134: 244114.

116 Perdew, J.P., Burke, K., and Ernzerhof, M. (1996). Generalized gradient approximation made simple. *Phys. Rev. Lett.* 77: 3865–3868.

117 Runge, E. and Gross, E.K.U. (1984). Density-functional theory for time-dependent systems. *Phys. Rev. Lett.* 52: 997–1000.

118 Petersilka, M., Gossmann, U.J., and Gross, E.K.U. (1996). Excitation energies from time-dependent density-functional theory. *Phys. Rev. Lett.* 76: 1212–1215.

119 Casida, M.E. (1995). Time-dependent density-functional response theory for molecules. In: *Recent Advances in Density Functional Methods* (ed. D.P. Chong), 155–192. Singapore: World Scientific.

120 Tamm, I. (1945). Relativistic interaction of elementary particles. *J. Phys. (USSR)* (9): 449.

121 Dancoff, S.M. (1950). Non-adiabatic meson theory of nuclear forces. *Phys. Rev.* 78: 382–385.

122 Kleinman, L. and Bylander, D.M. (1982). Efficacious form for model pseudopotentials. *Phys. Rev. Lett.* 48: 1425–1428.

123 Agostini, F., Min, S.K., Tavernelli, I., and Gossel, G.H. (2018). CTMQC. https://e-cam.readthedocs.io/en/latest/Quantum-Dynamics-Modules/modules/CTMQC/readme.html,

124 Longuet-Higgins, H.C. (1961). Some recent developments in the theory of molecular energy levels. In: *Advances in Spectroscopy*, vol. II (ed. H.W. Thompson), 429–472. Interscience Publishers, Ltd.

125 Berry, M.V. (1984). Quantal phase factors accompanying adiabatic changes. *Proc. R. Soc. Lond. A* 392: 45–57.

126 Yarkony, D.R. (2012). Nonadiabatic quantum chemistry: past, present and future. *Chem. Rev.* 112 (1): 481–498.

127 Engelman, R. (1972). *The Jahn-Teller Effect in Molecules and Crystals*. New York: Wiley-Interscience.

128 Ryabinkin, G. and Izmaylov, A.F. (2013). Geometric phase effects in dynamics near conical intersections: symmetry breaking and spatial localization. *Phys. Rev. Lett.* 111: 220406.

129 Ryabinkin, I.G., Joubert-Doriol, L., and Izmaylov, A.F. (2017). Geometric phase effects in non-adiabatic dynamics near conical intersections. *Acc. Chem. Res.* 50: 1785–1793.

130 Halász, G.J., Vibók, Á., Baer, R., and Baer, M. (2007). Conical intersections induced by the Renner effect in polyatomic molecules. *J. Phys. A Math. Theor.* 40 (15): F267–F272.

131 Baer, M. (1975). Adiabatic and diabatic representations for atom-molecule collisions: treatment of the collinear arrangement. *Chem. Phys. Lett.* 35: 112–118.

132 Baer, M. and Engelman, R. (1992). A study of the diabatic electronic representation within the Born–Oppenheimer approximation. *Mol. Phys.* 75: 293–303.

133 Domcke, W., Köppel, H., and Cederbaum, L. (1981). Spectroscopic effects of conical intersections of molecular potential energy surfaces. *Mol. Phys.* 43: 851–875.

134 Baer, M. (2002). Introduction to the theory of electronic non-adiabatic coupling terms in molecular systems. *Phys. Rep.* 358: 75–142.

135 Hunter, G. (1974). Conditional probability amplitude analysis of coupled harmonic oscillators. *Int. J. Quantum Chem.* 8: 413–420.

136 Berry, M.V. (1989). The quantum phase, five years after. In: *Geometric Phases in Physics* (eds. A. Shapere and F. Wilczek), 7–28. World Scientific.

18

Bohmian Approaches to Non-Adiabatic Molecular Dynamics

Guillermo Albareda[1] *and Ivano Tavernelli*[2]

[1] *Max Planck Institute for the Structure and Dynamics of Matter and Center for Free-Electron Laser Science, Luruper Chaussee 149, 22761 Hamburg, Germany*
[2] *IBM Research GmbH, Zurich Research Laboratory, Säumerstrasse 4, 8803 Rüschlikon, Switzerland*

Abstract

In this chapter we focus on Bohmian mechanics and on how this route to quantum mechanics offers an alternative to the so-called quantum wave packet methods. We will differentiate between approaches to non-adiabatic dynamics based on the Born-Huang decomposition of the molecular wave function and approaches that are formulated in the position-space representation of the full electron-nuclear Schrödinger equation. In the Born-Huang basis, time-dependent wave packets representing the nuclear motion are discretized into a set of Bohmian fluid elements and are followed in time by integrating coupled equations of motion that are solved in the Lagrangian picture of fluid motion according to the Hamilton-Jacobi equations. Techniques based on this framework have been applied to describe considerably large molecules, mainly thanks to their suitability to be coupled with well established electronic structure methods. Alternatively, in Bohmian approaches based on the position-space representation of the full molecular wave function, time-dependent electrons (nuclei) are described by waves that parametrically depend, via the Coulombic potential energy of the system, on nuclear (electronic) trajectories. These approaches allow to bypass the, typically necessary, computation of multiple Born-Oppenheimer potential-energy surfaces and non-adiabatic coupling terms and hence offer a new and attractive route to calculate molecular observables and time-correlation functions.

18.1 Introduction

Understanding the several formulations (and the connections among them) of a given mathematical theory not only helps to better appreciate its physical soundness but also allows new approaches to be conceived for addressing a particular physical problem. In quantum mechanics, while most practical problems are solved using the so-called orthodox or Copenhagen [1] interpretation, other well established formulations such as Feynman's path integrals have innumerable and very successful applications in, e.g., quantum statistics or quantum field theory [2].

In this chapter we focus on Bohmian mechanics [3] and on how this alternative formulation of quantum mechanics can lead to novel numerical approaches for addressing molecular dynamics problems. Bohmian mechanics was proposed by Louis de Broglie [4, 5] even before the Copenhagen explanation of quantum phenomena was established. In the 1950s, David Bohm [6–8] clarified the

Quantum Chemistry and Dynamics of Excited States: Methods and Applications,
First Edition. Edited by Leticia González and Roland Lindh.

meaning and applications of this explanation of quantum mechanics in terms of point-like particles guided by waves. Historically, Bohmian mechanics has been involved in many metaphysical disputes about the role of waves and particles when trying to provide a hierarchy between different quantum formulations. However, as far as one looks for a formalism that reproduces experimental results, all quantum theories (orthodox, Bohmian, Feynman, many-worlds, etc.) are perfectly valid. Hence, the relevant question that we address in this chapter is whether Bohmian mechanics can help in the numerical description of non-adiabatic quantum dynamics.

One of the most important developments in molecular dynamics, is the so-called ab-initio molecular dynamics method, which combines nuclear dynamics with forces obtained from electronic structure calculations. The Born–Oppenheimer potential energy surfaces (BOPESs) are the central concept for understanding ab-initio molecular dynamics. BOPESs have been mapped out with higher and higher accuracy for larger and larger molecular systems with accurate first principles electronic structure methods (such as density functional theory, DFT, or time-dependent density functional theory, TDDFT). Under the assumption that electrons adjust adiabatically to the slower motion of the nuclei, nuclear dynamics simulations have been carried out on top of single BOPESs, both assuming classical equations of motion or with more accurate quantum mechanical propagation schemes for small systems, sometimes with spectacular success in reproducing experiments [9].

Many challenging chemical processes, however, cannot be properly described with a single BOPES. The assumption that electrons adjust instantaneously to the motion of the nuclei becomes meaningless whenever electronic and nuclear motions occur on comparable time-scales [10]. Electronic (non-adiabatic) transitions between different BOPESs play, indeed, a pivotal role in numerous chemical processes, such as electron transfer in electrochemical reactions, ion-molecule reactions, or in proton-coupled electron transfer [11]. Similarly, electronic transitions are essential to asses the performance of single-molecule electronic devices [12].

To study these non-adiabatic processes it is necessary to go beyond the quasi-static view of the electron–nuclear interaction. Mixed quantum-classical approaches, where electrons are treated quantum mechanically and the nuclei are described with classical mechanics, have become particularly appealing because of the localized nature of the nuclei in many relevant scenarios. The interaction between classical and quantum degrees of freedom is usually addressed assuming a self-consistent field, i.e., nuclei evolve on top of a single effective potential energy surface defined as a weighted average of the involved adiabatic BOPESs. Branching of the nuclear wave packet into several BOPESs is not well described by these (Ehrenfest-like) approaches because the back-reaction between classical and quantum subsystems is described under mean-field assumptions [13–16]. See, e.g., a detailed derivation of the Ehrenfest approach in Chapter 15. Multi-configuration schemes, such as Tully's surface hopping, are in general required to account for bifurcation paths [17, 18] (see also a detailed discussion of the Surface Hopping approach in Chapter 16). Despite the undeniable success of these mixed approaches to describe many non-adiabatic phenomena, some limitations arise when quantum nuclear effects such as tunneling [19], decoherence [20] or interferences [21] occur.

Bohmian mechanics offers a trajectory-based scheme to describe quantum nuclear effects, and in this way represents an alternative to the so-called quantum wave packet methods [22, 23] (a detailed derivation of the basic equations of motion in the multi-configurational time-dependent Hartree method can be found, e.g., in Chapter 12). Since the pioneer work of Wyatt in 1999 [24], several schemes based on Bohmian mechanics have been proposed to describe molecular dynamics beyond the adiabatic regime. In this chapter we will differentiate between approaches to non-adiabatic dynamics based on the so-called Born–Huang (BH) decomposition of the molecular wave function and approaches that are formulated in the position-space representation of the

full electron–nuclear Schrödinger equation. In practice, all quantum dynamics methods are essentially formulated using a discrete (truncated) Hilbert space representation for the electronic degrees of freedom. In this picture, the Born–Oppenheimer approximation naturally emerges as classical nuclear dynamics on the electronic ground state BOPES [25], and non-adiabatic effects are introduced by including multiple electronic potential energy surfaces and non-adiabatic coupling terms (NACTs) [23]. An alternative to the BH picture, is to use the (real space) position representation for both electrons and nuclei. This allows to go beyond the BO picture without the need to explicitly calculate several BOPESs and NACTs [26]. This is an attractive feature from a computational point of view, as these quantities may be demanding to obtain from *ab-initio* electronic structure calculations.

This chapter is organized as follows. In Section 18.2 we provide a quick introduction to Bohmian mechanics including a hands-on oriented list of its essential mathematical ingredients. In Section 18.3 we introduce the time-dependent molecular Schrödinger equation as it is written in the full (electron–nuclear) position-space representation. In the same section, we will derive the equations of motion of electrons and nuclei as the position basis is replaced by the so-called Born–Huang basis set. Section 18.4 will be devoted to discussing Bohmian approaches that are derived starting from the Born–Huang representation of the molecular wave function. Alternatively, in Section 18.5 we will discuss Bohmian approaches that are based on the full position representation of the molecular wave function. A summary of the state-of-the-art of Bohmian approaches to non-adiabatic dynamics is provided in Section 18.6.

18.2 A Practical Overview of Bohmian Mechanics

Bohmian mechanics is presented in this section in a very practical way, i.e., paying attention only to those mathematical elements that will be strictly necessary later in practical implementations. The particularization of the formalism to molecular dynamics problems will depend on the choice of the basis, viz., Born–Huang or position-space, and hence will be detailed later in Sections 18.4 and 18.5 respectively.

18.2.1 The Postulates

We first present the theory of Bohmian mechanics through a small set of very short and simple working postulates. These Bohmian postulates are only valid for non-relativistic quantum systems, where the number of particles does not change with time. The generalization of these postulates for systems described by quantum field theory is not within the scope of this chapter. Furthermore, and for the sake of brevity, we will assume spinless particles and will not consider at this point any symmetrization postulate. A more general practical introduction to Bohmian mechanics can be found in Ref. [27].

First Postulate (dynamics of a quantum system): The dynamics of a non-relativistic quantum system of N spinless particles comprises a many-particle wave function $\Psi(\mathbf{r}, t)$, defined in the configuration space $\mathbf{r} = (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$ and time t, and a many-particle trajectory $\mathbf{r}(t) = (\mathbf{r}_1(t), \mathbf{r}_2(t), \ldots, \mathbf{r}_N(t))$ that evolves continuously under the guidance of the wave function. The wave function is a solution of the Schrödinger equation:

$$i\hbar \frac{\partial \Psi(\mathbf{r}, t)}{\partial t} = \left(\sum_{k=1}^{N} T_k + V(\mathbf{r}, t) \right) \Psi(\mathbf{r}, t) \tag{18.1}$$

where T_k is the kinetic energy operator of particle k and the potential $V(\mathbf{r}, t)$ includes all interactions in the system (internal and with an arbitrary external scalar potential). Each component $\mathbf{r}_k(t)$ of the trajectory is obtained by time-integrating the particle velocity $\mathbf{v}_k(t) = \mathbf{v}_k(\mathbf{r}(t), t)$ defined through the velocity field $\mathbf{v}_k(\mathbf{r}, t) = \mathbf{j}_k(\mathbf{r}, t)/|\Psi(\mathbf{r}, t)|^2$, where $\mathbf{j}_k(\mathbf{r}, t)$ is the kth particle current density.

Second Postulate (quantum equilibrium hypothesis): The initial position $\mathbf{r}(t_0)$ of the trajectory $\mathbf{r}(t)$ cannot be known with certainty, and it is randomly distributed according to the quantum probability density $|\Psi(\mathbf{r}, t_0)|^2$. This condition can be written mathematically as:

$$|\Psi(\mathbf{r}, t_0)|^2 = \lim_{M\to\infty} \frac{1}{M} \sum_{\alpha=1}^{M} \prod_{k=1}^{N} \delta(\mathbf{r}_k - \mathbf{r}_k^{\alpha}(t_0)) \tag{18.2}$$

Notice the presence of two indices, $\alpha = 1, \ldots, M$ denotes an infinite ensemble of trajectories accounting for the initial uncertainty and $k = 1, \ldots, N$ accounts for the total number N of particles. The initial velocity of the trajectory $\mathbf{r}^{\alpha}(t)$ is then determined by $\mathbf{v}_k^{\alpha}(t_0) = \mathbf{v}_k(\mathbf{r}^{\alpha}(t_0), t_0)$.

18.2.2 Computation of Bohmian Trajectories

In Bohmian mechanics, a quantum system is described by both a wave function and a particle position which describes a well-defined trajectory guided by the wave function. There are mainly two approaches to compute the dynamics of a system in Bohmian mechanics, viz., the analytic and the synthetic algorithms to which we will dedicate this section.

The basis of the analytic approaches consists in computing first the wave function and then obtaining the Bohmian trajectories from it (see Section 18.2.2.1). In a sense, the trajectories do not contribute to the structure of the algorithm, but are simply obtained by the equations in the formalism. While these algorithms do not add, in principle, any computational advantage, e.g., the trajectory computation is an additional step to integrating the Schrödinger equation, they can be easily implemented to obtain the trajectory dynamics which can be very useful to gain insights into the dynamics. On the other hand, they are at the foundation of the conditional wave function algorithms which will be discussed in Section 18.5. We note here that the conditional wave function concept is different from the conditional probability that appears in the exact factorization approach to molecular dynamics (see Chapter 17 for a detailed derivation of the exact factorization method).

Alternatively, in the synthetic algorithms, Bohmian trajectories play a key part in the algorithm to perform the computations, i.e., as the points where the wave function is evaluated (see Section 18.2.2.2). Thus, these algorithms require an extra step in formulating them, i.e., the solution of the quantum Hamilton–Jacobi equation. This "synthetic" approach to Bohmian mechanics was the one used by David Bohm in his formulation of the theory in the early 1950s [6, 7]. This approach allows to obtain the trajectories without first computing the wave function, and it is the source of many hydrodynamic algorithms, some of them discussed in Section 18.4. Furthermore, it sets the basis for extensions such as Bohmian mechanics with complex action which will be briefly described in Section 18.2.2.3.

18.2.2.1 Trajectories from the Schrödinger Equation

As in standard quantum mechanics, the time evolution of the wave function is given by the Schrödinger equation in Eq. (18.1). The solutions of the Schrödinger equation obey a continuity equation. From Eq. (18.1) it is easy to see that

$$\frac{\partial \rho(\mathbf{r}, t)}{\partial t} + \nabla \cdot \mathbf{J}(\mathbf{r}, t) = 0, \tag{18.3}$$

where we have defined a probability density and its associated current respectively as $\rho(\mathbf{r},t) = |\Psi(\mathbf{r},t)|^2$ and

$$\mathbf{J}(\mathbf{r},t) = \sum_k \mathbf{j}_k(\mathbf{r},t) = \sum_k \left[\frac{\hbar}{m_k}\mathrm{Im}(\Psi^*(\mathbf{r},t)\nabla_k\Psi(\mathbf{r},t))\right] \tag{18.4}$$

These densities lead to the definition of a velocity field,

$$\mathbf{v}(\mathbf{r},t) = \frac{\mathbf{J}(\mathbf{r},t)}{\rho(\mathbf{r},t)} = \sum_k \left[\frac{\hbar}{m_k}\mathrm{Im}\left(\frac{\nabla_k\Psi(\mathbf{r},t)}{\Psi(\mathbf{r},t)}\right)\right], \tag{18.5}$$

which provides a guidance law for an N-particle trajectory $\mathbf{r}^\alpha(t) = (\mathbf{r}_1^\alpha(t), \ldots, \mathbf{r}_N^\alpha(t))$:

$$\dot{\mathbf{r}}^\alpha(t) = \mathbf{v}(\mathbf{r}^\alpha(t),t). \tag{18.6}$$

The initial positions $\{\mathbf{r}^\alpha(0)\}$ of the trajectories $\{\mathbf{r}^\alpha(t)\}$ are distributed according to the quantum equilibrium hypothesis in Eq. (18.2). Then, the continuity equation in (18.3) ensures that the trajectories will be distributed following $\rho(\mathbf{r},t)$ at all later times. It is important to note that the velocity field associated with each particle is defined on the entire configuration space. Specifically, $\dot{\mathbf{r}}_k^\alpha(t) = \mathbf{v}_k(\mathbf{r}^\alpha(t),t)$, where $\mathbf{v}_k(\mathbf{r},t) = \mathbf{j}_k(\mathbf{r},t)/\rho(\mathbf{r},t)$. Thus the trajectory of each particle in the system experiences non-local effects through the positions of the remaining particles.

18.2.2.2 Trajectories from the Hamilton–Jacobi Equation

Two working equations are obtained by expressing the wave function in polar form, $\Psi(\mathbf{r},t) = \mathcal{A}(\mathbf{r},t)e^{iS(\mathbf{r},t)/\hbar}$, and then introducing it into the Schrödinger equation (18.1), viz.:

$$\frac{\partial\mathcal{A}^2}{\partial t} + \nabla\cdot\sum_k \mathcal{A}^2\frac{\nabla_k S}{m_k} = 0, \tag{18.7}$$

and

$$\frac{\partial S(\mathbf{r},t)}{\partial t} + \sum_k \frac{(\nabla_k S(\mathbf{r},t))^2}{2m_k} + V(\mathbf{r},t) + Q(\mathbf{r},t) = 0. \tag{18.8}$$

The first equation is equivalent to the continuity equation in (18.3), while Eq. (18.8) is known as the quantum Hamilton–Jacobi equation, analogous to its classical counterpart but with Q as an additional potential term (commonly known as quantum potential) defined as:

$$Q(\mathbf{r},t) = -\sum_k \frac{\hbar^2}{2m_k}\frac{\nabla_k^2\mathcal{A}(\mathbf{r},t)}{\mathcal{A}(\mathbf{r},t)}. \tag{18.9}$$

Equation (18.8) can then be used to describe an ensemble of trajectories defined by:

$$\dot{\mathbf{r}}^\alpha(t) = \sum_k \frac{\nabla_k S(\mathbf{r},t)}{m_k}\bigg|_{\mathbf{r}=\mathbf{r}^\alpha(t)}, \tag{18.10}$$

and initially sampled according to the quantum equilibrium hypothesis in Eq. (18.2). Notice that Eq. (18.10) is equivalent to Eq. (18.6).

By taking the limit $Q \to 0$ the (classical) Hamilton–Jacobi equation is recovered, from where classical trajectories would be obtained. Since Q accounts for the quantum (and non-local) behavior of the trajectories, it is named the quantum potential, and its magnitude gives an estimation of the deviation of quantum trajectories from their classical counterparts. Nevertheless, thinking of it as a classical potential can be misleading since it depends on the shape of the wave function (cf. Eq. (18.9)).

The numerical integration of the Hamilton–Jacobi equation is more convoluted than the Schrödinger equation. To begin with, the Hamilton–Jacobi equation is a nonlinear equation (with respect to the modulus of the wave function), and thus numerical instabilities are bound to appear more easily. Furthermore, the computation of the wave function (modulus and phase) in regions where the modulus is small (for instance, near wave function nodes) should be handled with special care because, depending on the implemented algorithm, the trajectories in those areas can become sparse. Algorithms such as the derivative propagation and trajectory stability methods [28] were proposed to avoid this kind of problem.

18.2.2.3 Trajectories from a Complex Action

As we may find in classical and semiclassical treatments [29], Bohmian mechanics can also be recast in a complex form in terms of a complex action, $\bar{S}$, and extended by analytic continuation to the complex plane. This complexification gives rise to alternative dynamical behaviors, which are specified by a complex-valued time-dependent quantum Hamilton–Jacobi equation,

$$\frac{\partial \bar{S}}{\partial t} + \frac{(\nabla \bar{S})^2}{2m} + V - \frac{i\hbar}{2m} \nabla^2 \bar{S} = 0, \tag{18.11}$$

where the last term on the left-hand side is the complex quantum potential. The relationship between $\bar{S}$ and the usual wave function is given by the transformation relation

$$\bar{S}(\mathbf{r}, t) = \frac{\hbar}{i} \ln \Psi(\mathbf{r}, t). \tag{18.12}$$

From this equation, one can now define a complex-valued local velocity vector field,

$$\bar{\mathbf{v}}(\mathbf{r}, t) = \frac{\nabla \bar{S}}{m}. \tag{18.13}$$

Taking this expression into account, the complex quantum potential can be expressed in terms of the first spatial derivative of the complex velocity:

$$\bar{Q}(\mathbf{r}, t) = -\frac{i\hbar}{2m} \nabla^2 \bar{S} = -\frac{i\hbar}{2} \nabla \bar{\mathbf{v}}. \tag{18.14}$$

That is, within this formulation in terms of a complex action, also known as *complex Bohmian mechanics*, there is a direct relationship between the quantum potential and the local velocity field, thus stressing the direct role of $\bar{Q}$ on the quantum dynamics.

Because $\bar{S}$ is in general a complex field, the only dynamics compatible with Eq. (18.13) has to be complex, which means that we cannot use the real variable, $\mathbf{r}$, but must use a complex one, $\mathbf{z}$, obtained by analytic continuation. This means that the corresponding complex trajectories are obtained after integration of the (complex) equation of motion $d\mathbf{z}/dt = \bar{\mathbf{v}}$.

A direct correspondence cannot be established between the trajectories obtained from this equation and the usual Bohmian trajectories in real space, since a one to one correspondence among them does not exist. Rather, each Bohmian trajectory is to be considered as the result of the crossing of the real axis, at subsequent times, of a continuous set of complex trajectories [30].

The complex version of Bohmian mechanics has been invoked as an alternative computational tool, the so-called *Bohmian mechanics with complex action*, developed since 2006 by Tannor and coworkers [31–37] from an earlier, independent derivation of Eq. (18.11) [38], with extensions to non-adiabatic molecular dynamics [39, 40]. Bohmian mechanics with complex action aims to obtain the wave function directly from the trajectories, as the *quantum trajectory method* [24] or the *derivative propagation methods* [41] do in real space. In other words, this complexified Bohmian mechanics is also an alternative synthetic method but in complex space [42].

18.2.3 Computation of Expectation Values

To see how the language of operators can be merged with a trajectory-based formulation of quantum mechanics, one can proceed as follows. The Hermitian operator $\hat{O}$ and the expectation value $\langle\hat{O}\rangle_\Psi$ can always be written in the position representation. Then, the mean value of this operator over the wave function $\Psi(\mathbf{r},t)$ is given by:

$$\langle\hat{O}\rangle_\Psi = \int_{-\infty}^{\infty} \Psi^*(\mathbf{r},t)O(\mathbf{r},i\hbar\nabla)\Psi(\mathbf{r},t)d\mathbf{r}. \tag{18.15}$$

Alternatively, the same mean value can be computed from Bohmian mechanics by defining a spatial average of a "local" magnitude $O_B(\mathbf{r},t)$ weighted by $\mathcal{A}^2(\mathbf{r},t)$:

$$\langle\hat{O}\rangle_\Psi = \int_{-\infty}^{\infty} \mathcal{A}^2(\mathbf{r},t)O_B(\mathbf{r},t)d\mathbf{r}. \tag{18.16}$$

In order to obtain the same result with Eqs. (18.15) and (18.16), one can easily identify the local mean value $O_B(\mathbf{r},t)$ with

$$O_B(\mathbf{r},t) = \mathrm{Re}\left[\frac{\Psi^*(\mathbf{r},t)O(\mathbf{r},i\hbar\nabla)\Psi(\mathbf{r},t)}{\Psi^*(\mathbf{r},t)\Psi(\mathbf{r},t)}\right]_{\Psi=\mathcal{A}e^{i\frac{S}{\hbar}}}, \tag{18.17}$$

where only the real part is taken into account because the mean value of Eq. (18.17) is real.

For practical purposes, expectation values are computed using Eq. (18.16) by means of a large $\alpha = 1,\ldots,M$ number of Bohmian trajectories with initial positions selected according to the quantum equilibrium hypothesis Eq. (18.2). The initial positions $\mathbf{r}^\alpha(t_0)$ of the trajectories are used to rewrite $\mathcal{A}^2(\mathbf{r},t)$ in Eq. (18.16) as:

$$\langle\hat{O}\rangle_\Psi = \lim_{M\to\infty}\frac{1}{M}\sum_{\alpha=1}^{M} O_B(\mathbf{r}^\alpha(t)). \tag{18.18}$$

By construction, in the limit $M\to\infty$, the value of Eq. (18.18) is identical to the value of Eq. (18.15) and Eq. (18.16).

18.3 The Born–Huang Picture of Molecular Dynamics

18.3.1 The Molecular Schrödinger Equation in Position Space

The starting point for the development of quantum dynamics in molecular systems is the time-dependent Schrödinger equation:

$$H_{mol}\Psi(\boldsymbol{r},\boldsymbol{R},t) = i\hbar\frac{\partial}{\partial t}\Psi(\boldsymbol{r},\boldsymbol{R},t), \tag{18.19}$$

where $\boldsymbol{R} = (\boldsymbol{R}_1,\boldsymbol{R}_2,\ldots,\boldsymbol{R}_{N_n})$ is the collective vector of the N_n nuclear positions in $\mathbb{R}^{3N_i}$ and $\boldsymbol{r} = (\boldsymbol{r}_1,\boldsymbol{r}_2,\ldots,\boldsymbol{r}_{N_{el}})$ the one for the N_{el} electrons. In Eq. (18.19), $\hat{H}_{mol}$ is the molecular Hamiltonian

$$\begin{aligned} H_{mol}(\boldsymbol{r},\boldsymbol{R}) &= -\sum_A\frac{\hbar^2}{2M_A}\nabla_A^2 - \sum_i\frac{\hbar^2}{2m}\nabla_i^2 + \sum_{i<j}\frac{1}{|\boldsymbol{r}_i-\boldsymbol{r}_j|} - \sum_{A,i}\frac{Z_A}{|\boldsymbol{R}_A-\boldsymbol{r}_i|} + \sum_{A<B}\frac{Z_AZ_B}{|\boldsymbol{R}_A-\boldsymbol{R}_B|} \\ &= -\sum_A\frac{\hbar^2}{2M_A}\nabla_A^2 + \mathcal{H}_{el}(\boldsymbol{r},\boldsymbol{R}) \end{aligned} \tag{18.20}$$

and $\Psi(\boldsymbol{r},\boldsymbol{R},t)$ is the total wave function of the nuclear (labelled A and B) and electronic (labelled i and j) degrees of freedom. Here γ and ζ will be used to label electronic states. Later, indices α and

β will be also used to label trajectories. In the following, atomic units will be used except for the reduced Planck constant $\hbar$ and the electronic mass m which will be retained for clarity.

In this first section, we will derive the equation of motion for the nuclear and electronic degrees of freedom using what is known as a trajectory-based approach. In this framework, the electrons are described at a quantum mechanical level, while the nuclear wave packet is discretized into an ensemble of points in phase space and then propagated along classical (or Bohmian quantum) trajectories that, as we will see, keep some essential quantum character including non-adiabatic effects.

The first step in the derivation of the equations of motion for the combined electron–nuclear dynamics is the definition of a suitable representation of the total system wave function. Depending on the particular choice of this expansion we can obtain different (approximated) solutions of the initial molecular Schrödinger equation (Eq. (18.19)). In the following we will restrict ourselves to the Born–Huang representation of the total molecular wave function, which can be used to derive trajectory-based non-adiabatic molecular dynamics solutions like Born–Oppenheimer dynamics and Bohmian dynamics.

The following derivation is partially inspired by the articles of J. C. Tully (Ref. [43]), the book of Marx and Hutter on *ab-initio* molecular dynamics [44], and Curchod et al. [45–48].

18.3.2 Schrödinger Equation in the Born–Huang Basis

The Born–Oppenheimer MD equations can be derived starting from the Born–Huang representation of the molecular wave function [49, 50]

$$\Psi(\boldsymbol{r},\boldsymbol{R},t) = \sum_{\gamma}^{\infty} \Omega_{\gamma}(\boldsymbol{R},t)\Phi_{\gamma}(\boldsymbol{r};\boldsymbol{R}) \; . \tag{18.21}$$

In this equation, $\{\Phi_{\gamma}(\boldsymbol{r};\boldsymbol{R})\}$ describes a complete set of orthonormal electronic wave function solutions of the time-independent Schrödinger equation

$$\hat{\mathcal{H}}_{el}(\boldsymbol{r};\boldsymbol{R})\Phi_{\gamma}(\boldsymbol{r};\boldsymbol{R}) = E^{el}_{\gamma}(\boldsymbol{R})\Phi_{\gamma}(\boldsymbol{r};\boldsymbol{R}) \tag{18.22}$$

with $\langle\Phi_{\zeta}|\Phi_{\gamma}\rangle = \delta_{\gamma\zeta}$ and where '$;\boldsymbol{R}$' denotes the parametric dependence of the electronic Schrödinger equation from the position of the atoms. Note that only the nuclear wave function depends explicitly on time, while $\hat{\mathcal{H}}_{el}(\boldsymbol{r};\boldsymbol{R})$ and $\Phi_{\gamma}(\boldsymbol{r};\boldsymbol{R})$ will only depend on t through the implicit time-dependence of $\boldsymbol{R}(t)$.

Inserting Eq. (18.22) into the time-dependent Schrödinger equation, Eq. (18.19), we obtain (after multiplying by $\Phi^{*}_{\zeta}(\boldsymbol{r};\boldsymbol{R})$ from the left-hand-side and integrating over the electronic degrees of freedom, $\boldsymbol{r}$)

$$i\hbar\frac{\partial}{\partial t}\Omega_{\zeta}(\boldsymbol{R},t) = \left[-\sum_{A}\frac{\hbar^2}{2M_A}\nabla^2_A + E^{el}_{\zeta}(\boldsymbol{R})\right]\Omega_{\zeta}(\boldsymbol{R},t) + \sum_{\gamma}\mathcal{F}_{\gamma\zeta}\Omega_{\gamma}(\boldsymbol{R},t) \tag{18.23}$$

The quantities $\mathcal{F}_{\gamma\zeta}(\boldsymbol{R})$

$$\begin{aligned}\mathcal{F}_{\gamma\zeta}(\boldsymbol{R}) = &\int d\boldsymbol{r}\; \Phi^{*}_{\zeta}(\boldsymbol{r};\boldsymbol{R})\left[-\sum_{A}\frac{\hbar^2}{2M_A}\nabla^2_A\right]\Phi_{\gamma}(\boldsymbol{r};\boldsymbol{R}) \\ &+\sum_{A}\frac{1}{M_A}\left\{\int d\boldsymbol{r}\; \Phi^{*}_{\zeta}(\boldsymbol{r};\boldsymbol{R})[-i\hbar\nabla_A]\Phi_{\gamma}(\boldsymbol{r};\boldsymbol{R})\right\}[-i\hbar\nabla_A]\end{aligned} \tag{18.24}$$

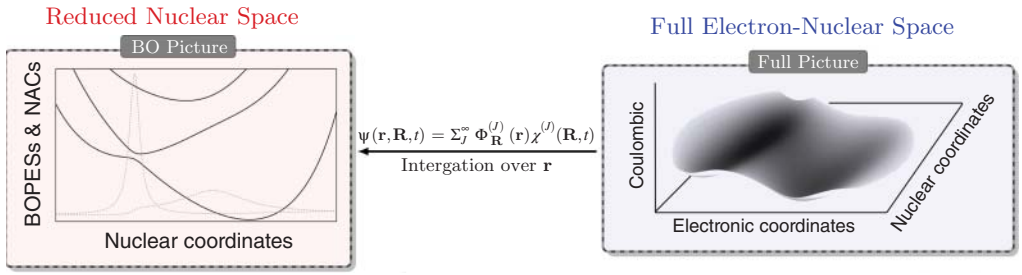


Figure 18.1 Schematic depiction of two different representations for the description of the correlated electron–nuclear motion: (Left panel) the Born–Huang expansion of the electron–nuclear wave function, and (right panel) the full configuration space picture. We distinguish between approaches where the nuclear dynamics occurs in a reduced nuclear subspace, on Born–Oppenheimer (BO) potential energy surfaces, and approaches where it occurs in the full electron–nuclear Hilbert space.

are the *non-adiabatic couplings* (NACs), with a contribution from the nuclear kinetic energy operator and a second from the momentum operator. In the most general case, the non-diagonal elements of $\mathcal{F}_{\gamma\zeta}(\boldsymbol{R})$ are non-zero and induce a coupling between different electronic states due to the motion of the nuclei. In fact, the last term in Eq. (18.23) brings amplitude $(\mathcal{F}_{\gamma\zeta}\Omega_\gamma(\boldsymbol{R},t))$ from the "electronic state with energy" $E_\zeta^{el}(\boldsymbol{R})$, solution of the electronic time-independent Schrödinger equation (18.22)), to the actual state ζ, with energy $E_\zeta^{el}(\boldsymbol{R})$. This interpretation of the nuclear wave function dynamics Eq. (18.23) is at the basis of almost all approaches to molecular dynamics including, e.g., the *surface hopping* description of non-adiabatic dynamics. See Figure 18.1 for a pictorial comparison between the position-space and the BH representations of the time-dependent molecular Schrödinger equation.

18.3.2.1 The Born–Oppenheimer Approximation: The Adiabatic Case

In the *Born–Oppenheimer approximation* only the diagonal terms, $\mathcal{F}_{\zeta\zeta}$, are retained

$$\mathcal{F}_{\zeta\zeta} = \int \Phi_\zeta^*(\boldsymbol{r};\boldsymbol{R}) \left[-\sum_A \frac{\hbar^2}{2M_A}\nabla_A^2 \right] \Phi_\zeta(\boldsymbol{r};\boldsymbol{R}) d\boldsymbol{r} \,, \tag{18.25}$$

which only induce a shift of the electronic potential energy surfaces $E_\zeta^{el}(\boldsymbol{R})$ felt by the nuclear wave functions (the second term of Eq. (18.24) is zero for $\gamma = \zeta$, when $\Phi_\gamma(\boldsymbol{r};\boldsymbol{R})$ are real).

In this approximation, the nuclei move in the potential of a single electronic state, the potential energy surface (PES) $E_\zeta^{el}(\boldsymbol{R})$, and the electronic (Eq. (18.22)) and nuclear (Eq. (18.23)) Schrödinger equations become completely decoupled. The term $\mathcal{F}_{\zeta\zeta}(\boldsymbol{R})$ is called the Born–Oppenheimer diagonal correction [51, 52] and, depending on the nuclear mass, induces an isotope-dependence [53–55] of the total energy, $E_\zeta^{el}(\boldsymbol{R}) + \mathcal{F}_{\zeta\zeta}(\boldsymbol{R})$. However, this term is usually small and is neglected in the so-called *Born–Oppenheimer adiabatic approximation* [56].

In order to derive a trajectory-based representation of the nuclear dynamics, we introduce the polar representation of the nuclear wave function $\Omega_\zeta(\boldsymbol{R},t)$ as we did in Section 18.2.2,

$$\Omega_\zeta(\boldsymbol{R},t) = \mathcal{A}_\zeta(\boldsymbol{R},t)\exp\left[\frac{i}{\hbar}S_\zeta(\boldsymbol{R},t)\right]. \tag{18.26}$$

Inserting this polar representation into Eq. (18.23) (with all $\mathcal{F}_{\gamma\zeta}(\boldsymbol{R}) = 0$) and separating the real and the imaginary parts, we obtain

$$\frac{\partial S_\zeta}{\partial t} = \frac{\hbar^2}{2m}\sum_A \frac{1}{M_A}\frac{\nabla_A^2\mathcal{A}_\zeta}{\mathcal{A}_\zeta} - \frac{1}{2}\sum_A \frac{1}{M_A}(\nabla_A S_\zeta)^2 - E_\zeta^{el} \tag{18.27}$$

$$\frac{\partial \mathcal{A}_\zeta}{\partial t} = -\sum_A \frac{1}{M_A} \nabla_A \mathcal{A}_\zeta \cdot \nabla_A S_\zeta - \frac{1}{2} \sum_A \frac{1}{M_A} \mathcal{A}_\zeta \nabla_A^2 S_\zeta \tag{18.28}$$

where all dependences S, A and E_ζ^{el} from the electronic and nuclear coordinates are omitted to simplify the notation.

Taking the classical limit $\hbar \to 0$ [57] in Eq. (18.27), we obtain a Hamilton–Jacobi equation for the action function $S_\zeta(\boldsymbol{R}, t)$

$$\frac{\partial S_\zeta}{\partial t} = -\frac{1}{2} \sum_A \frac{1}{M_A} (\nabla_A S_\zeta)^2 - E_\zeta^{el} \,, \tag{18.29}$$

which correspond to a classical point-particle time evolution of the nuclei, and from Eq. (18.28) a continuity equation [58] for the propagation of the amplitude on the adiabatic state of interest, $d/dt(\int d\boldsymbol{R}\ |\Omega_\zeta(\boldsymbol{R}, t)|^2) = 0$. We observe that the potential acting on the nuclei is obtained from a static expectation value of the electronic Hamiltonian computed for the time-independent state $\Phi_\zeta(\boldsymbol{r}; \boldsymbol{R})$ solution of the electronic Schrödinger equation Eq. (18.22).

Using the relation, $\nabla_A S_\zeta|_{\boldsymbol{R}(t)} = \boldsymbol{P}_\zeta^A(t)$, we obtain a Newton-like equation of motion for the "classical" nuclei:

$$M_A \ddot{\boldsymbol{R}}_A(t) = -\nabla_A E_\zeta^{el}(\boldsymbol{R}(t)) \,. \tag{18.30}$$

In summary, the Born–Oppenheimer molecular dynamics (BO-MD) equations can be described by the following system of coupled equations

$$\hat{\mathcal{H}}_{el}(\boldsymbol{r}; \boldsymbol{R}(t)) \Phi_\zeta(\boldsymbol{r}; \boldsymbol{R}(t)) = E_\zeta^{el}(\boldsymbol{R}(t)) \Phi_\zeta(\boldsymbol{r}; \boldsymbol{R}(t)) \tag{18.31}$$

$$M_A \ddot{\boldsymbol{R}}_A(t) = -\nabla_A E_\zeta^{el}(\boldsymbol{R}(t)) = -\underset{min \Phi_\zeta}{\nabla_A} \langle \Phi_\zeta | \hat{\mathcal{H}}_{el} | \Phi_\zeta \rangle \,, \tag{18.32}$$

where only the second one describes an explicit time evolution. The electronic energies and the forces acting on the nuclei are computed *statically* solving Eq. (18.31) *on-the-fly* at each new position sampled along the trajectory $\boldsymbol{R}(t)$. Note that in BO-MD there is no explicit time-dependence of the electronic degrees of freedom. It is important to further stress that, due to the assumption that $\mathcal{F}_{\gamma\zeta} = 0$, the BO-MD always evolves on a single electronic PES, even in the case in which the system approaches regions of strong coupling between electronic and nuclear degrees of freedom. In practice, the state of interest is the ground state for which the adiabatic separation from all other states (excited states) holds in most (non-metallic) cases.

The combination of BO-MD with DFT for the on-the-fly calculation of the electronic structure properties (energies and forces) at each MD step is straightforward and can be found in many textbooks (see for instance [44]). Using the Hohenberg–Kohn theorem one first maps the electronic structure problem from the wave function space into the density space and then, within the Kohn–Sham formulation of DFT, the electronic ground state energy functional, $E_0[\rho(\boldsymbol{r}; \boldsymbol{R})]$, and its gradients are computed.

18.3.2.2 Non-Adiabatic Dynamics

The adiabatic approximation breaks down when the electronic states get close in energy, which especially occurs when the dynamics is initiated in one of the electronic excited states of the system. This is the usual situation encountered in a *pump–probe* experiment, where an initial pulse is exciting the system while a second one is monitoring its time-dependent relaxation towards the ground state (or a stable excited state).

The starting point is the time-dependent Schrödinger equation for the molecular system Eq. (18.23) that we rewrite as

$$i\hbar\frac{\partial\Omega_\zeta(\boldsymbol{R},t)}{\partial t} = -\sum_A \frac{\hbar^2}{2M_A}\nabla_A^2\Omega_\zeta(\boldsymbol{R},t) + E_\zeta^{el}(\boldsymbol{R})\Omega_\zeta(\boldsymbol{R},t) + \sum_{A\gamma}\frac{\hbar^2}{2M_A}D_{\gamma\zeta}^A(\boldsymbol{R})\Omega_\gamma(\boldsymbol{R},t) - \sum_{A,\gamma\neq\zeta}\frac{\hbar^2}{M_A}\boldsymbol{d}_{\gamma\zeta}^A(\boldsymbol{R})\cdot\nabla_A\Omega_\gamma(\boldsymbol{R},t) \tag{18.33}$$

where

$$\boldsymbol{d}_{\gamma\zeta}^A(\boldsymbol{R}) = \int \Phi_\zeta^*(\boldsymbol{r};\boldsymbol{R})\nabla_A\Phi_\gamma(\boldsymbol{r};\boldsymbol{R})d\boldsymbol{r} \tag{18.34}$$

are the first-order non-adiabatic couplings (or non-adiabatic coupling vectors), and

$$D_{\gamma\zeta}^A(\boldsymbol{R}) = -\int \Phi_\zeta^*(\boldsymbol{r};\boldsymbol{R})\nabla_A^2\Phi_\gamma(\boldsymbol{r};\boldsymbol{R})d\boldsymbol{r} \tag{18.35}$$

are the second-order coupling elements.

18.4 BH-Based Approaches

As was shown in Section 18.3, electrons and nuclei are treated asymmetrically in the Born–Huang picture of molecular dynamics. Specifically, electrons are described through a time-independent Schrödinger equation that is parametrically dependent on the nuclear degrees of freedom that are, at the same time, recast into an effective time-dependent Schrödinger equation. Starting from this picture, in Bohmian approaches to non-adiabatic dynamics, the nuclear motion is represented by time-dependent wave packets that are discretized into a set of Bohmian fluid elements, these are propagated in time by integrating coupled equations of motion that are solved in the Lagrangian picture of fluid motion according to the Hamilton–Jacobi equations (see Section 18.2.2).

18.4.1 The Non-Adiabatic Bohmian Dynamics Equations (NABDY)

Using the polar representation for the nuclear wave function Eq. (18.26) in Eq. (18.33) we obtain, after separating real and imaginary parts,

$$\begin{aligned}-\frac{\partial S_\zeta(\boldsymbol{R},t)}{\partial t} &= \sum_A \frac{1}{2M_A}(\nabla_A S_\zeta(\boldsymbol{R},t))^2 + E_\zeta^{el}(\boldsymbol{R}) - \sum_A \frac{\hbar^2}{2M_A}\frac{\nabla_A^2\mathcal{A}_\zeta(\boldsymbol{R},t)}{\mathcal{A}_\zeta(\boldsymbol{R},t)} \\ &+ \sum_{A\gamma}\frac{\hbar^2}{2M_A}D_{\gamma\zeta}^A(\boldsymbol{R})\frac{\mathcal{A}_\gamma(\boldsymbol{R},t)}{\mathcal{A}_\zeta(\boldsymbol{R},t)}\Re[e^{i\phi_{\gamma\zeta}(\boldsymbol{R},t)}] - \sum_{A,\gamma\neq\zeta}\frac{\hbar^2}{M_A}\boldsymbol{d}_{\gamma\zeta}^A(\boldsymbol{R}) \\ &\cdot\frac{\nabla_A\mathcal{A}_\gamma(\boldsymbol{R},t)}{\mathcal{A}_\zeta(\boldsymbol{R},t)}\Re[e^{i\phi_{\gamma\zeta}(\boldsymbol{R},t)}] \\ &+ \sum_{A,\gamma\neq\zeta}\frac{\hbar}{M_A}\frac{\mathcal{A}_\gamma(\boldsymbol{R},t)}{\mathcal{A}_\zeta(\boldsymbol{R},t)}\boldsymbol{d}_{\gamma\zeta}^A(\boldsymbol{R})\cdot\nabla_A S_\gamma(\boldsymbol{R},t)\Im[e^{i\phi_{\gamma\zeta}(\boldsymbol{R},t)}]\end{aligned} \tag{18.36}$$

and

$$\begin{aligned}\hbar\frac{\partial\mathcal{A}_\zeta(\boldsymbol{R},t)}{\partial t} =& -\sum_A\frac{\hbar}{M_A}\nabla_A\mathcal{A}_\zeta(\boldsymbol{R},t)\cdot\nabla_A S_\zeta(\boldsymbol{R},t)-\sum_A\frac{\hbar}{2M_A}\mathcal{A}_\zeta(\boldsymbol{R},t)\nabla_A^2 S_\zeta(\boldsymbol{R},t)\\ &+\sum_{A\gamma}\frac{\hbar^2}{2M_A}D^A_{\gamma\zeta}(\boldsymbol{R})\mathcal{A}_\gamma(\boldsymbol{R},t)\Im[e^{i\phi_{\gamma\zeta}(\boldsymbol{R},t)}]-\sum_{A,\gamma\neq\zeta}\frac{\hbar^2}{M_A}\boldsymbol{d}^A_{\gamma\zeta}(\boldsymbol{R})\\ &\cdot\nabla_A\mathcal{A}_\gamma(\boldsymbol{R},t)\Im[e^{i\phi_{\gamma\zeta}(\boldsymbol{R},t)}]\\ &-\sum_{A,\gamma\neq\zeta}\frac{\hbar}{M_A}\mathcal{A}_\gamma(\boldsymbol{R},t)\boldsymbol{d}^A_{\gamma\zeta}(\boldsymbol{R})\cdot\nabla_A S_\gamma(\boldsymbol{R},t)\Re[e^{i\phi_{\gamma\zeta}(\boldsymbol{R},t)}],\end{aligned}\tag{18.37}$$

where both $S_\zeta(\boldsymbol{R},t)$ and $\mathcal{A}_\zeta(\boldsymbol{R},t)$ are real fields and $\phi_{\gamma\zeta}(\boldsymbol{R},t)=\frac{1}{\hbar}(S_\gamma(\boldsymbol{R},t)-S_\zeta(\boldsymbol{R},t))$. Equations (18.36) and (18.37) correspond to the *exact* Schrödinger equation for a nuclear wave function evolving in the potential of the different electronic surfaces determined by the time-independent Schrödinger equation. In comparison with the single surface case or with the diabatic formulation, time evolution of phases and amplitudes involves first- and second-order coupling elements that mix contributions from other potential energy surfaces. In particular, transfer of amplitude from one PES to another becomes possible thanks to the coupling terms, which however, make the solution of the set of Eqs. (18.36) and (18.37) more involved. In the dynamics that emerges, the first equation is the equivalent of the classical Hamilton–Jacobi (HJ) equation (first two terms) for the action $S(\boldsymbol{R},t)$, augmented with two additional parts of quantum nature of order $\hbar$ and $\hbar^2$. The third term is the quantum potential $\mathcal{Q}_\zeta(\boldsymbol{R},t)$ describing all quantum effects on a single PES and introducing non-locality [59], and the fourth to sixth terms constitute the non-adiabatic quantum potential $\sum_\gamma\mathcal{D}_{\gamma\zeta}(\boldsymbol{R},t)$ describing interstate contributions.

After applying the gradient with respect to the nucleus B on both sides of Eq. (18.36) and moving to the Lagrangian frame [48], we obtain a Newton-like equation of motion (using the HJ definition of the momenta $\nabla_B S_\zeta(\boldsymbol{R},t)|_{\boldsymbol{R}(t)}=\boldsymbol{P}^B_\zeta(t)$)

$$M_B\frac{d^2\boldsymbol{R}_B(t)}{(dt^\zeta)^2}=-\nabla_B\left[E^{el}_\zeta(\boldsymbol{R}(t))+\mathcal{Q}_\zeta(\boldsymbol{R}(t),t)+\sum_\gamma\mathcal{D}_{\gamma\zeta}(\boldsymbol{R}(t),t)\right]\tag{18.38}$$

describing the time evolution (trajectory) of the $\boldsymbol{R}_B$ components of a fluid element with collective variable $\boldsymbol{R}(t)$ ($d/dt^\zeta=\partial/\partial t+\sum_A\nabla_A S_\zeta(\boldsymbol{R},t)/M_A\cdot\nabla_A$).

In the same moving frame, the time-evolution of the amplitudes becomes

$$\begin{aligned}\hbar\frac{d}{dt}\mathcal{A}_\zeta(\boldsymbol{R}(t),t)=&-\frac{\hbar}{2}\sum_A M_A^{-1}\mathcal{A}_\zeta(\boldsymbol{R}(t),t)\nabla_A^2 S_\zeta(\boldsymbol{R}(t))\\ &+\sum_\gamma H_{\gamma\zeta}(\boldsymbol{R}(t))\mathcal{A}_\gamma(\boldsymbol{R}(t),t)\Im\left[e^{(\frac{i}{\hbar}S_{\gamma-\zeta}(\boldsymbol{R}(t)))}\right]\\ &+\frac{\hbar^2}{2m}\sum_{A\gamma}\frac{1}{M_A}D^A_{\gamma\zeta}(\boldsymbol{R}(t))\mathcal{A}_\gamma(\boldsymbol{R}(t),t)\Im\left[e^{(\frac{i}{\hbar}S_{\gamma-\zeta}(\boldsymbol{R}(t)))}\right]\\ &-\hbar^2\sum_{A,\gamma\neq\zeta}\frac{1}{M_A}\mathbf{d}^A_{\gamma\zeta}\nabla_A\mathcal{A}_\gamma(\boldsymbol{R}(t),t)\Im\left[e^{(\frac{i}{\hbar}S_{\gamma-\zeta}(\boldsymbol{R}(t)))}\right]\\ &-\hbar\sum_{A,\gamma\neq\zeta}\frac{1}{M_A}\mathbf{d}^A_{\gamma\zeta}\mathcal{A}_\gamma(\boldsymbol{R}(t),t)\nabla_A S_\gamma(\boldsymbol{R}(t))\Re\left[e^{(\frac{i}{\hbar}S_{\gamma-\zeta}(\boldsymbol{R}(t)))}\right].\end{aligned}\tag{18.39}$$

Note that the difficulties associated with the non-crossing rule between fluid elements in configuration space [59] becomes less severe as the dimensions of the system increase (the most difficult case being the one-dimensional dynamics).

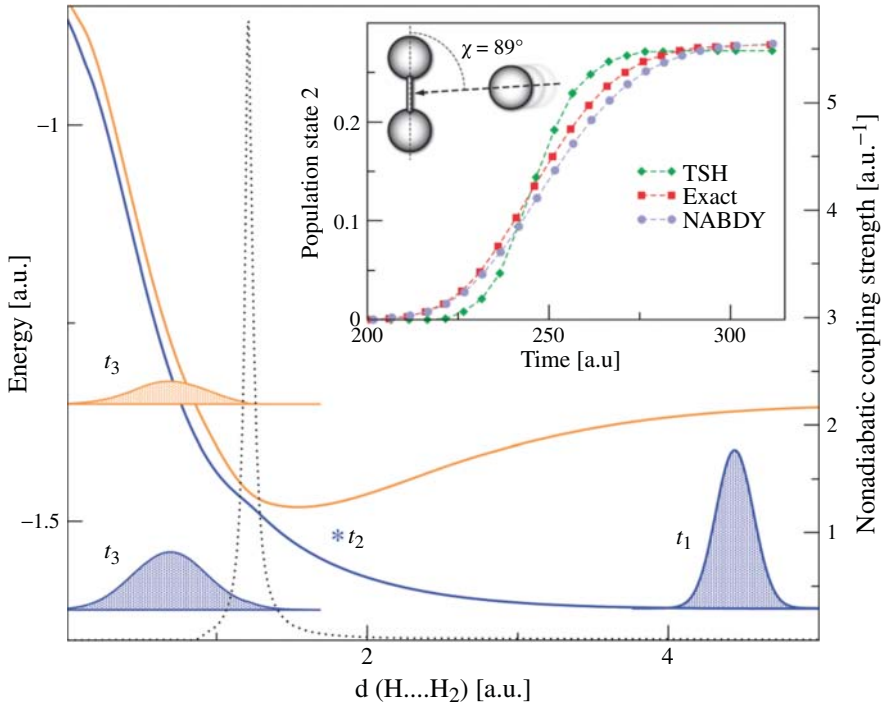


Figure 18.2 NABDY applied to the collision of H with H2 ($\chi = 89^{\circ}$ and d(H-H) = 1.4 a.u., see the inset). An initial Gaussian wave packet (wp) is prepared on the ground state ($t_1 = 0$ a.u.) with an initial momentum $k = 75$ a.u. Shown is the probability density of the nuclear wp obtained with 352 trajectories at the initial time (t_1) and after the region of coupling ($t_3 = 300$ a.u.). The displacement of the resulting wps in the vertical direction is arbitrary. Blue: wp on state 1; orange: wp on state 2; black dotted line: non-adiabatic coupling strength. The inset shows the time evolution of the population transfer obtained using the different schemes (TSH: 3112 trajectories).

The most severe limitation in the implementation of NABDY is related to the calculation of the derivatives in the configuration space such as that of the amplitudes, $\nabla_A \mathcal{A}_\zeta(\boldsymbol{R}(t), t)$, and of the phases $\nabla_A S_\zeta(\boldsymbol{R}(t), t)$. In fact, as the size of the system increases, it becomes very impractical to perform such derivatives on a real grid of dimensions $3N_n - 6$. In Figure 18.2 we show an application of NABDY in the study of the collision of a fast moving hydrogen atom with a standing hydrogen molecule [60]. The initial conditions are chosen such that the system evolves trough a geometry that lies close to a conical intersection between the ground and the first excited state (equilateral structure) leading to a "thermal" or "kinetic" induced excitation of the system.

18.4.2 Implementation in Molecular Dynamics: The Adiabatic Case

The NABDY approach can be adapted to perform quantum dynamics in the high-dimensional, unconstrained, phase space of large molecular systems. In this case, the main challenge lies in the calculation of the phase space derivatives of the quantum potential and nuclear amplitudes, which are best performed analytically [61]. To this end, we introduce a decomposition of the nuclear wave packet into a sum of Gaussian functions, which are then propagated in time using a set of differential equations derived from quantum hydrodynamics equations. Within this numerical approximation, the phase space derivatives used in the calculation of the quantum potential

can be performed analytically, thus providing a solution to the instability problem related to the discontinuities of the quantum potential.

The applications of this approach are, however, restricted to the adiabatic case, which well illustrate the theoretical and numerical challenges associated to this type of dynamics and its level of accuracy without introducing additional complications related to non-adiabaticity. Further numerical investigations are needed to extend the method to the non-adiabatic case.

The main step consists of the representation of the molecular amplitude wave function with the product

$$A(\boldsymbol{R}_1(t), \boldsymbol{R}_2(t), \boldsymbol{R}_3(t), \dots, \boldsymbol{R}_{N_n}(t), t) = \Phi_1(\boldsymbol{R}_1(t), t)\Phi_2(\boldsymbol{R}_2(t), t)\Phi_3(\boldsymbol{R}_3(t), t) \dots \Phi_{N_n}(\boldsymbol{R}_{N_n}(t), t) \tag{18.40}$$

where $\Phi_A(\boldsymbol{R}_A(t), t)$ is given by a sum of Gaussians centered at $\boldsymbol{R}_A^\alpha(t) \in \mathbb{R}^3$

$$\begin{aligned}\Phi_A(\mathbf{R}_\mathbf{A}(t), t) &= \sum_{\alpha=1}^{M} \tilde{\phi}_\alpha^\alpha(\boldsymbol{R}_A(t) - \boldsymbol{R}_A^\alpha(t); a_A^\alpha(t), \sigma_A^\alpha(t)) \\ &= \sum_{\alpha=1}^{M} \phi_A^\alpha(\boldsymbol{R}_A(t); \boldsymbol{R}_A^\alpha(t), a_A^\alpha(t), \sigma_A^\alpha(t))\end{aligned} \tag{18.41}$$

and

$$\phi_A^\alpha(\boldsymbol{R}_A(t); \boldsymbol{R}_A^\alpha(t), a_A^\alpha(t), \sigma_A^\alpha(t)) = \frac{a_A^\alpha(t)}{\mathcal{N}_A^\alpha(t)} e^{-\frac{(\boldsymbol{R}_A(t) - \boldsymbol{R}_A^\alpha(t))^2}{2(\sigma_A^\alpha(t))^2}} , \tag{18.42}$$

where $\mathcal{N}_A^\alpha(t) = (2\pi)^{3/2}(\sigma_A^\alpha(t))^3$. The solution of Eq. (18.39) with the product amplitude given in Eqs. (18.41) and (18.42) is still computationally very impractical. Therefore, we tentatively propose a dynamics in which the centers of the Gaussians follow the time evolution of the trajectories of Eq. (18.38) while the amplitudes of the Gaussians obey Eq. (18.39). This approach can be formalized using a (coarse-grained) configuration-space representation of the amplitude [62, 63]

$$A(\boldsymbol{R}, t) = \prod_A \sum_{\alpha=1}^{M} A(\boldsymbol{R}_A^\alpha, t)\delta(\boldsymbol{R}_A - \boldsymbol{R}_A^\alpha) , \tag{18.43}$$

where the $\boldsymbol{R}_A^\alpha$ are uniformly distributed points in the configuration space with associated amplitude $A(\boldsymbol{R}_A^\alpha, t)$ and evolve according to Eq. (18.38). The residual, explicit dynamics of $A(\boldsymbol{R}, t)$ is described by Eq. (18.39). For computational purposes, we introduce a broadening of the Dirac-delta function and work with the representation

$$A(\boldsymbol{R}, t) = \prod_A \sum_{i=1}^{M} A(\boldsymbol{R}_A^\alpha, t) g_A^\alpha(\boldsymbol{R}_A - \boldsymbol{R}_A^\alpha), \tag{18.44}$$

where $g_A^\alpha(\boldsymbol{R}_A - \boldsymbol{R}_A^\alpha)$ stays for a Gaussian of the form given in Eq. (18.42) (that gives the Dirac-delta in the limit of σ_A^α going to zero), $A(\boldsymbol{R}_A^\alpha, t) = a_A^\alpha(t)$, and the centers of the Gaussians, $\boldsymbol{R}_A^\alpha$, evolve – once more – along the trajectories in Eq. (18.38). The Gaussians appearing in Eq. (18.44) have nothing in common with the Gaussian wave packets used in previous works (see for instance reference [64]) and are not individual solutions of the original time-dependent Schrödinger equation, but they mainly serve as support for the amplitude dynamics given in Eq (18.39).

This approach has been applied to the study of the tunneling-enhanced proton transfer (PT) process in a protonated di-ammonia complex: $2(NH_3)H^+$ (Fig. IV B) [61]. The two nitrogen atoms are kept fixed at a distance such that PT cannot occur using a classical description of the nuclear

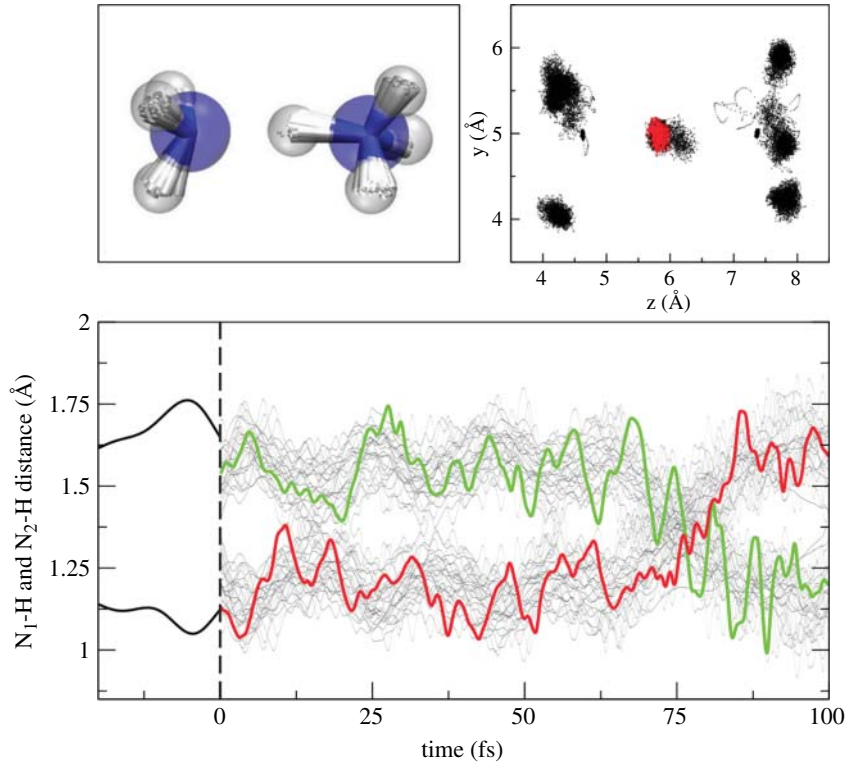


Figure 18.3 Upper left panel: ball-and-stick representation of the $(NH_3)_2\ H^+$ system. Thirty fluid elements (FEs) per atom are used in the quantum dynamics (small spheres). Larger atom-centered vdW spheres are used to wrap up all atomic FEs (blue: nitrogen, white hydrogen). Upper right panel: collection of all FE centers collected along the entire simulation. Lower panel: time-series of the two *NH* distances with the shared proton; ($t < 0$): dynamics with classical nuclei; ($t > 0$): dynamics with quantum trajectories. Two characteristic paths are shown in color to highlight the transition.

dynamics (in this case DFT-based Born–Oppenheimer dynamics using the PBE [65] functional). However, when the dynamics is switched to quantum using the ABDY approach described above (see also Ref. [61]) the PT becomes possible in the fs timescale (Fig. 18.3). This example shows how a suitable choice for the representation of the nuclear wave packet (in this case using an atom-centered Gaussian basis set) can lead to a stable propagation of the Bohmian quantum trajectories.

18.4.3 The Approximate Quantum Potential Approach

As pointed out in the previous section, the numerical implementation of the quantum potential poses challenges due to the fact that $Q(\mathbf{R}(t), t)$ is singular at the nodes of the nuclear wave packet, which makes the dynamics extremely sensitive to the accuracy of the corresponding quantum forces near the nodes.

In Section 18.4.2, we presented an expansion of the nuclear wave packet as a product of Gaussians centered at the position of the atoms. The reconstructed wave packet (Eq. (18.44)), although approximate, has the very interesting properties of being analytically differentiable in all space, including in the vicinity of the nodes.

A similar strategy was also developed in the group of S. Garashchuk [66] using a semi-empirical implementation of Bohmian quantum trajectories based on an approximate quantum potential (AQP). In this approach, the non-classical quantum momentum

$$\mathcal{P}(\boldsymbol{R},t) = \frac{\nabla A(\boldsymbol{R},t)}{A(\boldsymbol{R},t)} \tag{18.45}$$

is approximated by a linear combination of a small set of N_b basis functions $\boldsymbol{f}(\boldsymbol{R}) = (f_1(\boldsymbol{R}), f_2(\boldsymbol{R}), \ldots)$, so that

$$\tilde{\mathcal{P}}(\boldsymbol{R},t) = \boldsymbol{c}(t) \cdot \boldsymbol{f}(\boldsymbol{R}) \tag{18.46}$$

and

$$\tilde{\mathcal{Q}}(\boldsymbol{R},t) = -\frac{\tilde{\mathcal{P}}^2(\boldsymbol{R},t) + \nabla^2\tilde{\mathcal{P}}(\boldsymbol{R},t)}{2M}. \tag{18.47}$$

where M is the mass. Within this approximation, all gradients needed for the propagation of the quantum trajectory

$$\frac{d\boldsymbol{R}}{dt} = \frac{\boldsymbol{P}(\boldsymbol{R},t)}{M} \tag{18.48}$$

$$\frac{d\boldsymbol{P}(\boldsymbol{R},t)}{dt} = -\nabla V(\boldsymbol{R}) - \nabla\mathcal{Q}(\tilde{\boldsymbol{R}},t)(\boldsymbol{R}) = -\frac{d\boldsymbol{P}_{cl}(\boldsymbol{R},t)}{dt} - \frac{d\tilde{\mathcal{P}}(\boldsymbol{R},t)}{dt} \tag{18.49}$$

are obtained analytically [66]. The value of the expansion coefficients $\boldsymbol{c}(t)$ are obtained from the minimization of the function

$$I = \langle(\mathcal{P}(\tilde{\boldsymbol{R}},t) - \frac{\nabla A(\boldsymbol{R},t)}{A(\boldsymbol{R},t)})^2\rangle \tag{18.50}$$

by setting

$$\nabla_c I = 0\ . \tag{18.51}$$

Note that using the evolution equation for the trajectories given in Eqs. (18.48) and (18.49), the weights associated with the volume element $\delta\boldsymbol{R}$ of each trajectory

$$w(\boldsymbol{R},t) = A^2(\boldsymbol{R},t)\ \delta\boldsymbol{R}(t) \tag{18.52}$$

remains constant in time, $\frac{dw(\boldsymbol{R},t)}{dt} = 0$.

The equations of motion (Eqs. (18.48) and (18.49)) for the quantum trajectories are then solved using the AQP approximation for the spatial derivatives of the momenta

$$\frac{d\boldsymbol{P}(\boldsymbol{R},t)}{dt} = -\nabla V(\boldsymbol{R}) + \frac{1}{2M}(2\tilde{\mathcal{P}}(\boldsymbol{R},t)\nabla_{\boldsymbol{R}} + \nabla^2_{\boldsymbol{R}})\ \tilde{\mathcal{P}}(\boldsymbol{R},t) \tag{18.53}$$

$$\frac{d\tilde{\mathcal{P}}(\boldsymbol{R},t)}{dt} = -\frac{1}{2M}(2\tilde{\mathcal{P}}(\boldsymbol{R},t)\nabla_{\boldsymbol{R}} + \nabla^2_{\boldsymbol{R}})\ \boldsymbol{P}(\boldsymbol{R},t)\ . \tag{18.54}$$

The $\tilde{\mathcal{P}}(\boldsymbol{R},t)$ becomes therefore a trajectory-based variable sampled along the dynamics. Note that the function $\mathcal{P}(\boldsymbol{R},t)$ derived from Eq. (18.45) can be compared at each time step with the approximated function $\tilde{\mathcal{P}}(\boldsymbol{R},t)$, giving a measure of the error associated to the AQP approximation.

A further simplification of this scheme can be obtained using the linearized quantum force approximation (LQF) obtained introducing an expansion of $\tilde{\mathcal{P}}$ in a linear basis [67]. This approach is exact for Gaussian wave packets and, in general, is capable of describing leading quantum effects, such as the wave packet branching, tunneling, and zero point energy (ZPE). Advantages and disadvantages of the LQF approach are carefully discussed in Refs. [68–71]

18.5 Non-BH Approaches

A number of Bohmian approaches to electron–nuclear coupled dynamics have been also derived without relying on the Born–Huang basis set, i.e., by considering the molecular Schrödinger equation in Eq. (18.19) as the starting point. These approaches allow to bypass the, typically necessary, computation of multiple BOPESs and NACTs and hence offer a new and attractive route to calculate observables and time correlation functions without relying on the widely used Born–Oppenheimer picture. Furthermore, the calculation of quantum potentials, which is a major bottleneck for methods based on the Born–Huang basis, is also avoided by these approaches.

Here we will focus on two approaches, viz., the Conditional Wave function approach [26, 72–74], recently developed by the authors of this chapter, and the so-called time-dependent Monte Carlo approach to molecular dynamics developed by I. P. Christov [75–80]. Other relevant techniques have been derived, e.g., in the groups of C. Meier [81–84] or O. Prezhdo [85, 86].

18.5.1 The Conditional Wave Function Approach

The conditional wave function (CWF) approach is an exact decomposition and recasting of the unitary time-evolution of a closed quantum system, that yields a set of coupled, non-Hermitian, equations of motion [72]. The CWF approach allows one to describe the evolution of arbitrary subsets of the degrees of freedom in a system, on a formally exact level. In addition, this alternative formulation of the many-body quantum dynamics problem allows novel approximate schemes to be developed [73, 87] providing a completely new perspective to deal with the long-standing problems of non-adiabatic dynamics in complex interacting systems. Remarkably, the resulting propagation scheme does not require the computation of the quantum potential, in this manner overcoming a bottleneck in quantum trajectory-based approaches. We note here that the conditional wave function concept is different from the conditional probability that arises in the context of the exact factorization approach to molecular dynamics (see Chapter 17 for a detailed derivation of the exact factorization method).

The CWF approach can be developed starting from the full molecular Schrödinger equation in Eq. (18.19). The total Hamiltonian for the molecular system in Eq. (18.20) can be rewritten as

$$H_{mol} = T_e(\mathbf{r}) + T_n(\mathbf{R}) + W(\mathbf{r}, \mathbf{R}, t), \tag{18.55}$$

where the kinetic energy operators for each species of electrons and nuclei have been respectively defined as $T_e(\mathbf{r}) = -\sum_i \frac{\hbar^2}{2m}\nabla_i^2$ and $T_n(\mathbf{R}) = -\sum_A \frac{\hbar^2}{2M_A}\nabla_A^2$, and the full Coulombic energy of the system is gathered in $W(\mathbf{r}, \mathbf{R}, t) = \sum_{i<j}\frac{1}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} - \sum_{A,i}\frac{Z_A}{|\boldsymbol{R}_A - \boldsymbol{r}_i|} + \sum_{A<B}\frac{Z_A Z_B}{|\boldsymbol{R}_A - \boldsymbol{R}_B|}$.

The molecular wave function, $\Psi(\mathbf{r}, \mathbf{R}, t)$, can be exactly decomposed in terms of the CWFs of either of the two subsystems as:

$$\psi_e^\alpha(\mathbf{r}, t) := \int d\mathbf{R}\,\delta(\mathbf{R}^\alpha(t) - \mathbf{R})\Psi(\mathbf{r}, \mathbf{R}, t), \tag{18.56}$$

$$\psi_n^\alpha(\mathbf{R}, t) := \int d\mathbf{r}\,\delta(\mathbf{r}^\alpha(t) - \mathbf{r})\Psi(\mathbf{r}, \mathbf{R}, t). \tag{18.57}$$

Using these definitions in Eq. (18.57) and Eq. (18.58), one can show that the CWFs, $\psi_e^\alpha(t)$ and $\psi_n^\alpha(t)$, obey the following equations of motion [72]:

$$i\hbar\frac{d\psi_e^\alpha}{dt} = (T_e(\mathbf{r}) + W(\mathbf{r},\mathbf{R}^\alpha,t) + \eta_n^\alpha(\mathbf{r},t))\psi_e^\alpha, \tag{18.58}$$

$$i\hbar\frac{d\psi_n^\alpha}{dt} = (T_i(\mathbf{R}) + W(\mathbf{r}^\alpha,\mathbf{R},t) + \eta_e^\alpha(\mathbf{R},t))\psi_n^\alpha, \tag{18.59}$$

where we have suppressed the explicit time-dependence of the coordinates, i.e., $\{\mathbf{r}^\alpha,\mathbf{R}^\alpha\} \equiv \{\mathbf{r}^\alpha(t),\mathbf{R}^\alpha(t)\}$. The complex potentials $\eta_n^\alpha(\mathbf{r},t)$ and $\eta_e^\alpha(\mathbf{R},t)$ are functionals of the full wave function, and are respectively defined as [26, 72]:

$$\eta_n^\alpha(\mathbf{r},t) = T_n\Psi(\mathbf{r},\mathbf{R},t)|_{\mathbf{R}^\alpha(t)} + i\hbar\nabla_{\mathbf{R}}\Psi(\mathbf{r},\mathbf{R},t)|_{\mathbf{R}^\alpha(t)}\cdot\dot{\mathbf{R}}^\alpha(t), \tag{18.60}$$

$$\eta_e^\alpha(\mathbf{R},t) = T_e\Psi(\mathbf{r},\mathbf{R},t)|_{\mathbf{r}^\alpha(t)} + i\hbar\nabla_{\mathbf{r}}\Psi(\mathbf{r},\mathbf{R},t)|_{\mathbf{r}^\alpha(t)}\cdot\dot{\mathbf{r}}^\alpha(t). \tag{18.61}$$

The conditional wave functions, (18.56) and (18.57), represent *slices* of the full wave function taken along the degrees of freedom of the two disjoint subsets (see Figure 18.4). Each individual conditional wave function constitutes an open quantum system, whose time-evolution is non-unitary, due to the complex potentials.

While not required in principle, in practice it is useful to propagate both the nuclear and electronic conditional wave functions, Eqs. (18.58) and (18.59), to compute the quantum trajectories via conditional velocity fields defined as

$$\dot{\mathbf{r}}_i^\alpha(t) = \frac{\hbar}{m}\,\mathrm{Im}\left[\frac{\nabla_i\psi_e^\alpha(\mathbf{r},t)}{\psi_e^\alpha(\mathbf{r},t)}\right]_{\mathbf{r}^\alpha(t)}, \tag{18.62}$$

$$\dot{\mathbf{R}}_A^\alpha(t) = \frac{\hbar}{M_A}\,\mathrm{Im}\left[\frac{\nabla_A\psi_n^\alpha(\mathbf{R},t)}{\psi_n^\alpha(\mathbf{R}^\alpha,t)}\right]_{\mathbf{R}^\alpha(t)}. \tag{18.63}$$

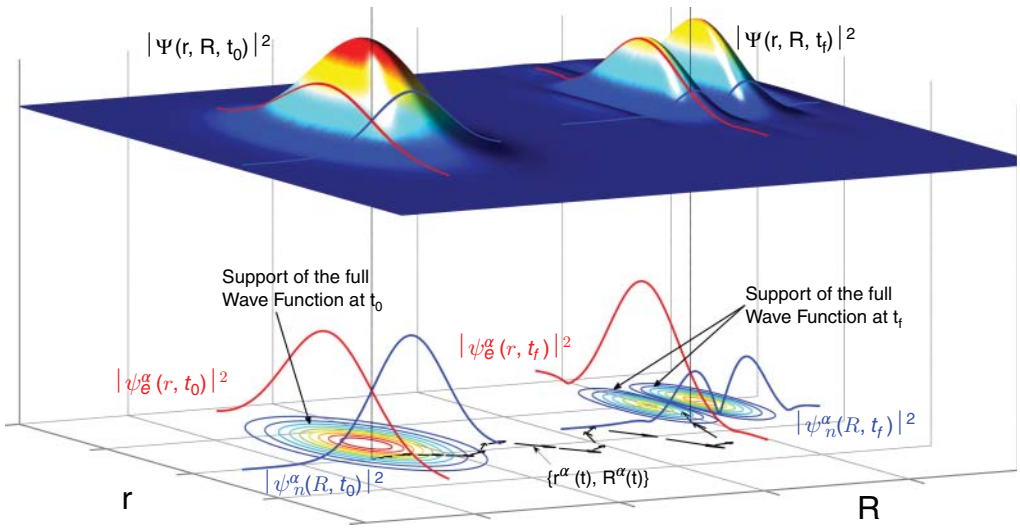


Figure 18.4 Schematic representation of the CWF approach to molecular dynamics for a two-dimensional system. The full nuclear probability-density $|\Psi(r,R,t)|^2$ is plotted at two different times t_0 and t_f, together with a pair of conditional amplitudes $|\psi_e^\alpha(r,t)|^2$ (in red) and $|\psi_i^\alpha(R,t)|^2$ (in blue) for a particular trajectory $\{r_\alpha(t),R_\alpha(t)\}$. Black arrows denote the velocity field $\{\dot{r}_\alpha(t),\dot{R}_\alpha(t)\}$, and contour plots of the full nuclear wave function are also shown for clarity.

In this way the reconstruction of the full wave function is avoided at the expense of solving the number of equations of motion twice [72]. Remarkably, the resulting propagation scheme, namely Eqs. (18.58) and (18.59) together with the trajectories in (18.71) and (18.72), does not require the computation of the quantum potential.

18.5.1.1 Hermitian Conditional Wave Function Approach

An approximate solution to Eqs. (18.58) and (18.59) can be formulated [72] by expanding the complex functionals around the conditional coordinates, and then truncating them such that $\eta_n^\alpha(\mathbf{r},t) = f(\mathbf{R}^\alpha,t)$ and $\eta_e^\alpha(\mathbf{R},t) = g(\mathbf{r}^\alpha,t)$. In this limit, these potentials only engender a pure time-dependent phase that can be omitted, as the conditional velocity fields are invariant under such global phase transformations [72]. The resulting propagation scheme is thus restored to a Hermitian form:

$$i\hbar\frac{d\psi_e^\alpha}{dt} = (T_e(\mathbf{r}) + W(\mathbf{r},\mathbf{R}^\alpha,t))\psi_e^\alpha, \tag{18.64}$$

$$i\hbar\frac{d\psi_n^\alpha}{dt} = (T_n(\mathbf{R}) + W(\mathbf{r}^\alpha,\mathbf{R},t))\psi_n^\alpha, \tag{18.65}$$

with the velocity fields defined as in Eqs. (18.71) and (18.72).

This approximate version of the CWF formalism is referred to as the Hermitian-CWF approach [72]. The Hermitian-CWF propagation scheme recasts the full quantum time-propagator as a set of independent single-species propagators, which is clearly a major simplification of the full problem. Hence, this form of the conditional decomposition allows one to circumvent the problem of storing and propagating the full many-body wave function, whose size scales exponentially with the number of degrees of freedom. Furthermore, the propagation of Eqs. (18.64) and (18.65) does not entail integrals over the electronic degrees of freedom and hence it is expected to be of particular interest in scenarios where several BOPESs and external electromagnetic fields are involved.

The Hermitian-CWF has been numerically tested using the so-called Shin–Metiu model system, which consists of three positively charged nuclei (ions) and a single electron in one dimension [88]. This model is very flexible and, based on the parameter regime chosen, can give rise to a number of challenging situations where electron–nuclear correlations play a crucial role in the dynamics. The Shin–Metiu model prescribes two ions that are fixed at a distance $L = 19.0a_0$, and the third ion and the electron are free to move in one dimension along the line joining the fixed ions (see Figure 18.5). The total Hamiltonian for the system can be written as:

$$H(r,R) = -\frac{1}{2m}\frac{\partial^2}{\partial r^2} - \frac{1}{2M}\frac{\partial^2}{\partial R^2} + W(r,R), \tag{18.66}$$

where m is the electron mass, and M is the proton mass. The coordinates of the electron and the mobile ion are measured from the center of the two fixed ions, and are labeled r and R, respectively. The full electron–nuclear potential reads:

$$W(r,R) = \frac{1}{|\frac{L}{2}-R|} + \frac{1}{|\frac{L}{2}+R|} - \frac{\mathrm{erf}\left(\frac{|R-r|}{R_f}\right)}{|R-r|} - \frac{\mathrm{erf}\left(\frac{|r-\frac{L}{2}|}{R_r}\right)}{|r-\frac{L}{2}|} - \frac{\mathrm{erf}\left(\frac{|r+\frac{L}{2}|}{R_l}\right)}{|r+\frac{L}{2}|}, \tag{18.67}$$

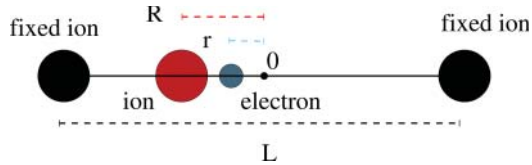


Figure 18.5 Schematic representation of the Shin–Metiu model [88]. Two ions are fixed (black) and a third one (red) and an electron (blue) are free to move in one dimension.

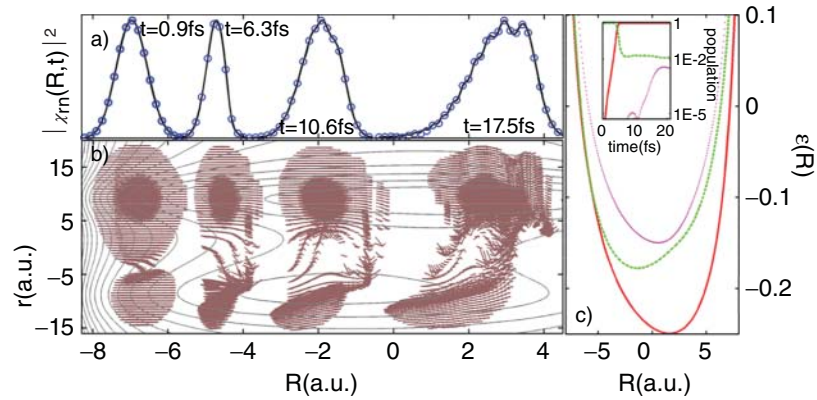


Figure 18.6 (a) Exact (black solid-line) and approximated (blue circles) nuclear probability densities at four different times. (b) Arrows refer to the (two-dimensional) velocity field computed from the approximated conditional wave functions. The gray contour lines represent the corresponding electron–nuclear two-dimensional potential energy surface. (c) First (red), second (green) and third (magenta) BOPESs involved in the non-adiabatic process. In the inset: adiabatic populations as a function of time computed from the exact solution.

where erf() represents the error function. In Ref. [72] the Shin–Metiu model system was studied under the parameter regime defined by: $R_f = 7a_0$, $R_l = 4.4a_0$ and $R_r = 3.1a_0$. In this way, the first BOPES, $\epsilon_{BO}^{(1)}$, is strongly coupled to the second BOPES, $\epsilon_{BO}^{(2)}$, within an extended region defined by $R < -4a_0$. In addition, there is a moderate coupling between the second BOPES, $\epsilon_{BO}^{(2)}$, and the third BOPES, $\epsilon_{BO}^{(3)}$ for $R > 2a_0$ (see Figure 18.6(c)). The coupling to the rest of the BOPESs is negligible. In Ref. [72], the system was initially excited to $\epsilon_{BO}^{(2)}$ and the initial nuclear wave function was defined to be a Gaussian wave packet with $\sigma = 1/\sqrt{2.85}$, centered at $R = -7.0a_0$. Starting with this initial state, one first samples its probability density with trajectories and then propagate Eq. (18.71) together with the Hermitian equations of motion in Eq. (18.64). In Fig.18.6.a snapshots at different times of the nuclear probability density are shown for the exact calculation (black solid line) and for the approximated solution (blue circles). Showing an excellent agreement, the Hermitian-CWF scheme is demonstrated to capture not only the conspicuous electronic transition between $\epsilon_{BO}^{(2)}$ and $\epsilon_{BO}^{(1)}$, but also the interferences originating at later times from contributions of higher adiabatic populations (see the rise of the population of $\epsilon_{BO}^{(3)}$ in the inset on Figure 18.6(c)). For this exactly solvable model system, the zero order approximation is able to accurately reproduce complex non-adiabatic dynamics (see Figure 18.6) with quantum nuclear effects included [72]. Furthermore, the use of Bohmian trajectories adds interpretative value to the method and provides a numerically stable algorithm to avoid the calculation of the unstable quantum potential.

18.5.2 The Interacting Conditional Wave Function Approach

Using a stochastic wave function ansatz that is based on a set of interacting single-particle Hermitian conditional wave functions, the recently developed Interacting-CWF propagation scheme allows to go beyond the Hermitian-CWF described in Section 18.5.1.1 while avoiding the computation of the non-local complex potentials, $\eta_n^\alpha(\mathbf{r}, t)$ and $\eta_e^\alpha(\mathbf{R}, t)$.

In Ref. [74] the following ansatz was considered for the full many-body wave function:

$$\Psi(\mathbf{r},\mathbf{R},t)=\sum_{\alpha=1}^{M}C_{\alpha}(t)\psi_{e}^{\alpha}(\mathbf{r},t)\psi_{n}^{\alpha}(\mathbf{R},t). \tag{18.68}$$

The basis functions in this sum are chosen to be Hermitian-CWFs, and the upper limit of the sum, M, refers to the total number of stochastically sampled trajectories (we will show below that these can be kept to a very low number, making the present scheme computationally very efficient). Including interactions between the trajectories in the ensemble corrects the Hermitian-CWF evolution, through the set of complex time-dependent coefficients, $\mathbf{C}(t)=\{C_1(t),...,C_M(t)\}$. The time evolution of these coefficients, is obtained by inserting Eq. (18.68) into Eq. (18.1),

$$i\hbar\mathbb{M}\dot{\mathbf{C}}(t)=(\mathbb{W}-\mathbb{W}_1-\mathbb{W}_2)\mathbf{C}(t), \tag{18.69}$$

where the matrix elements of $\mathbb{M}$, $\mathbb{W}$, $\mathbb{W}_1$, and $\mathbb{W}_2$ are:

$$\mathrm{M}_{\alpha,\alpha'}=\int d\mathbf{r}\psi_{e}^{\alpha'*}\psi_{e}^{\alpha}\int d\mathbf{R}\psi_{n}^{\alpha'*}\psi_{n}^{\alpha} \tag{18.70a}$$

$$\mathrm{W}_{\alpha,\alpha'}=\int d\mathbf{r}d\mathbf{R}\psi_{e}^{\alpha'*}\psi_{e}^{\alpha}\psi_{n}^{\alpha'*}\psi_{n}^{\alpha}W(\mathbf{r},\mathbf{R}) \tag{18.70b}$$

$$\mathrm{W}_{e}^{\alpha,\alpha'}=\int d\mathbf{r}\psi_{e}^{\alpha'*}\psi_{e}^{\alpha}W(\mathbf{r},\mathbf{R}^{\alpha})\int d\mathbf{R}\psi_{n}^{\alpha'*}\psi_{n}^{\alpha}, \tag{18.70c}$$

$$\mathrm{W}_{n}^{\alpha,\alpha'}=\int d\mathbf{r}\psi_{e}^{\alpha'*}\psi_{e}^{\alpha}\int d\mathbf{R}\psi_{n}^{\alpha'*}\psi_{n}^{\alpha}W(\mathbf{r}^{\alpha},\mathbf{R}). \tag{18.70d}$$

Obtaining these matrix elements is straightforward and, except for (18.71b), they can be easily calculated from independent single species integrals. Evaluating the matrix elements of $\mathbb{W}$, in principle, requires the reconstruction of the full (ansatz) wave function. This does not restrict the use of the method to cases where the potential energy $W(\mathbf{r},\mathbf{R})$ can be fit to a sum-of-products form, as in the multi-configurational time-dependent Hartree method [22] for example, but it does pose a potential numerical challenge in the case of a large trajectory ensemble (see a detailed description of this problem within the multi-configurational time-dependent Hartree method in, e.g., Chapter 12).

Once the coefficients $\mathbf{C}(t)$ are known, the velocity fields $\{\dot{\mathbf{r}}^{\alpha},\dot{\mathbf{R}}^{\alpha}\}$ are then constructed according to the exact expressions for each subsystem:

$$\dot{\mathbf{r}}_{i}^{\alpha}(t)=\frac{\hbar}{m}\,\mathrm{Im}\left[\frac{\sum_{\alpha}C_{\alpha}(t)\psi_{n}^{\alpha}(\mathbf{R}^{\alpha},t)(\nabla_{i}\psi_{e}^{\alpha}(\mathbf{r},t))|_{\mathbf{r}^{\alpha}(t)}}{\sum_{\alpha}C_{\alpha}(t)\psi_{e}^{\alpha}(\mathbf{r}^{\alpha},t)\psi_{n}^{\alpha}(\mathbf{R}^{\alpha},t)}\right], \tag{18.71}$$

$$\dot{\mathbf{R}}_{A}^{\alpha}(t)=\frac{\hbar}{M_A}\,\mathrm{Im}\left[\frac{\sum_{\alpha}C_{\alpha}(t)\psi_{e}^{\alpha}(\mathbf{r}^{\alpha},t)(\nabla_{A}\psi_{n}^{\alpha}(\mathbf{R},t))|_{\mathbf{R}^{\alpha}(t)}}{\sum_{\alpha}C_{\alpha}(t)\psi_{e}^{\alpha}(\mathbf{r}^{\alpha},t)\psi_{n}^{\alpha}(\mathbf{R}^{\alpha},t)}\right]. \tag{18.72}$$

The Interacting-CWF method, described above, does not require the electronic BOPES or NACs as input, or for time propagation. This feature is potentially quite advantageous for treating processes that involve many quantum states or continua, as in light-induced dynamics or surface-scattering phenomena.

In Ref. [74] the Shin–Metiu model (see Figure 18.8) system was studied using the Interacting-CWF for the parameter regime defined by: $R_f=5a_0$, $R_l=4a_0$ and $R_r=3.1a_0$. Under this condition, the ground BOPES, $\epsilon_{BO}^{(1)}$, is strongly coupled to the first excited adiabatic state, $\epsilon_{BO}^{(2)}$, around $R_{ac}=-2a_0$ (see Figure 18.7). The coupling to the rest of the BOPESs is negligible. The system

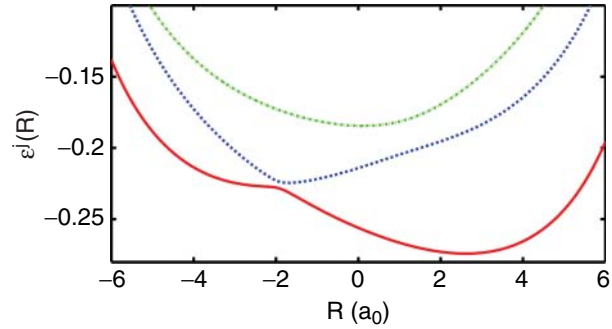


Figure 18.7 BOPESs for the Shin–Metiu model system. Ground state, $\epsilon^{(1)}$, first excited state, $\epsilon^{(2)}$, and second excited state, $\epsilon^{(3)}$, BOPESs are shown respectively as a red solid line, blue dashed line, and green dotted line.

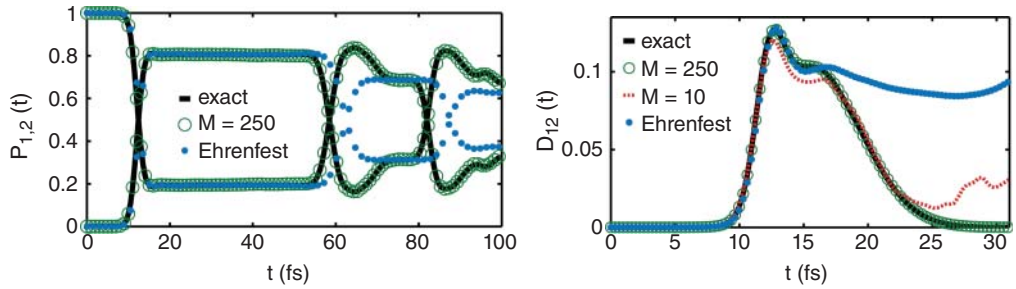


Figure 18.8 Dynamics in the Shin–Metiu model for photo-induced proton-coupled electron transfer [74]. Exact results (black line), Interacting-CWF with a number of trajectories $M = 250$ (open green circles) and Ehrenfest mean field theory (blue dots). Left panel: Evolution of the BO state populations $P_1(t)$ and $P_2(t)$. Right panel: Time-dependent decoherence indicator $D_{12}(t)$; Interacting-CWF results with 10 trajectories (red dashed line).

was supposed to be initially uncorrelated, as if prepared by a short laser pulse, in the first excited electronic state, $\epsilon^{(2)}_{BO}$, while the initial nuclear wave function is a Gaussian wave packet, with $\sigma = 1/\sqrt{2.85}$, has been supposed to be centered on the equilibrium geometry of the ground BO state, at $R = -4.0a_0$. The time-step used for integrating the TDSE is 2.4×10^{-3}fs (or $0.1a_0$). The fourth-order Runge–Kutta algorithm was used to propagate the CWF equations of motion and the corresponding trajectories. Propagating Eq. (18.70) requires the matrix inverse of $\mathbb{M}$ in Eq. (18.71a). However, this matrix can become ill-conditioned when different products of CWFs strongly overlap, or if the basis becomes overdetermined, for example. In Ref. [74] the Moore–Penrose pseudo-inversion method was used to ameliorate this numerical complication. Simulation results for the above described model system show that the Interacting-CWF method captures a quantitatively accurate physical picture, while using a number of trajectories that is orders of magnitude lower than the corresponding mean-field simulation (see Figure 18.8). Furthermore there is minimal cross-talk between trajectories, which makes the algorithm computationally efficient in massively parallel architectures. The degree of computational efficiency offered by this approach creates the possibility to treat dynamics in molecular and extended quantum systems with unprecedented accuracy without the need to pre-compute the BOPESs or NACTs, while providing access to all observables relevant for describing non-equilibrium dynamical phenomena.

18.5.3 Time-Dependent Quantum Monte Carlo

I. P. Christov has also presented an ab-initio method to solve quantum many-body problems of molecular dynamics where both electronic and nuclear degrees of freedom are represented by ensembles of Bohmian trajectories. In [89], the guiding waves are solutions of a set of approximated Schrödinger equations evaluated along electronic and nuclear trajectories. The quantum non-locality is incorporated into the model through effective potentials which are efficiently calculated by Monte Carlo integration. Unlike other many-body methods based on density functional calculations of the electronic structure, this approach uses explicit Coulomb potentials instead of parametrized exchange-correlation potentials.

The so-called Time-Dependent Quantum Monte Carlo (TDQMC) can be derive by starting with a single configuration ansatz (single Hartree product) for the total wave function where the nuclear and electronic coordinates are separated, $\Psi(\mathbf{r},\mathbf{R},t)=\chi(\mathbf{R},t)\Phi(\mathbf{r},t)$, where in addition we assume full factorization of the many-body wave functions for the nuclei $\chi(\mathbf{R},t)$ and for the electrons $\Phi(\mathbf{r},t)$, i.e.,:

$$\Psi(\mathbf{r},\mathbf{R},t)=\prod_{A=1}^{N_n}\chi_A(\mathbf{R}_A,t)\prod_{i=1}^{N_e}\Phi_i(\mathbf{r}_i,t). \tag{18.73}$$

In Eq. (18.74) it is assumed that the separate nuclear and electronic wave functions are normalized to unity at every instant of time. Following the standard procedure of Hartree theory one can derive a set of coupled nonlinear integro-differential equations for the nuclear and electronic wave functions,

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\chi_A(\mathbf{R}_A,t)=\Bigg[&\frac{-\hbar^2}{2M_A}\nabla_A^2+\sum_{B\neq A}\int d\mathbf{R}_B V_{n-n}(\mathbf{R}_A-\mathbf{R}_B)|\chi_B(\mathbf{R}_B,t)|^2 \\ &+\sum_i\int d\mathbf{r}_i V_{n-e}(\mathbf{R}_A-\mathbf{r}_i)|\Phi_i(\mathbf{r}_i,t)|^2+V_{ext}(\mathbf{R_A},t)\Bigg]\chi_A(\mathbf{R}_A,t) \end{aligned} \tag{18.74}$$

and

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\Phi_i(\mathbf{r}_i,t)=\Bigg[&\frac{-\hbar^2}{2m}\nabla_i^2+\sum_{j\neq i}\int d\mathbf{r}_j V_{e-e}(\mathbf{r}_i-\mathbf{r}_j)|\Phi_j(\mathbf{r}_j,t)|^2 \\ &+\sum_A\int d\mathbf{R}_A V_{e-n}(\mathbf{r}_i-\mathbf{R}_A)|\chi_A(\mathbf{R}_A,t)|^2+V_{ext}(\mathbf{r_i},t)\Bigg]\Phi_i(\mathbf{r}_i,t), \end{aligned} \tag{18.75}$$

where in Eqs. (18.74) and (18.75) $A,B=1,...,N_n$ and $i,j=1,...,N_e$. Terms which do not depend on $\mathbf{R}_A$ in Eq. (18.75) and on $\mathbf{r}_i$ in Eq. (18.76), respectively, have been omitted because these terms do not influence the motion of the corresponding Bohmian trajectories. Note that Eqs. (18.74) and (18.75) differ from the standard self-consistent field equations in that here we use one-body wave functions.

It is known, however, that the single determinant ansatz (Hartree) approximation and the resulting mean-field equations Eqs. (18.74) and (18.75) disregard important local and non-local quantum correlation effects. One approach to overcome these difficulties is to use a multi-configuration ansatz where multiple wave functions for the electron degree are used, which leads to multi-configuration time-dependent self-consistent field theory [22] (a detailed derivation of the basic equations of motion in the multi-configurational time-dependent Hartree method can be found, e.g., in Chapter 12). The TDQMC methodology assigns a separate set of wave functions

and trajectories to each physical particle where the wave functions guide the Bohmian trajectories for both the electron and nuclear degrees according to:

$$\dot{\mathbf{r}}_i^\alpha(t) = \frac{1}{m}\,\mathrm{Im}\left[\frac{\nabla_i \Phi(\mathbf{r},t)}{\Phi(\mathbf{r},t)}\right]_{\mathbf{r}_i^\alpha(t)}, \tag{18.76}$$

$$\dot{\mathbf{R}}_A^\alpha(t) = \frac{1}{M_A}\,\mathrm{Im}\left[\frac{\nabla_A \chi(\mathbf{R},t)}{\chi(\mathbf{R},t)}\right]_{\mathbf{R}_A^\alpha(t)}. \tag{18.77}$$

An important quantum correlation effect which is neglected in Hartree approximation is related to a specific quantum non-locality that arises due to the dependence of many-body wave functions in Eq. (18.74) on the coordinates in $3K + N$ dimensional configuration space. This non-locality is evidence of an interaction between different points in configuration space which represent the momentary coordinates of different replicas of the quantum system. One simple and efficient way to account for these quantum effects is to formally represent the particle densities in (18.74) and (18.75) by smoothed interpolation with, e.g., Gaussian kernels that are centered at the positions of the Bohmian trajectories (kernel density estimation). For the electrons one then has

$$|\Phi_i(\mathbf{r}_i,t)|^2 = \sum_{\alpha=1}^{M} \frac{1}{z_i^\alpha} \exp\left[-\frac{|\mathbf{r}_i - \mathbf{r}_i^\alpha(t)|^2}{\sigma_i^\alpha(\mathbf{r}_i^\alpha,t)^2}\right], \tag{18.78}$$

and for the nuclei,

$$|\chi_A(\mathbf{R}_A,t)|^2 = \sum_{\alpha=1}^{M} \frac{1}{Z_A^\alpha} \exp\left[-\frac{|\mathbf{R}_A - \mathbf{R}_A^\alpha(t)|^2}{\Sigma_A^\alpha(\mathbf{R}_A^\alpha,t)^2}\right], \tag{18.79}$$

where M is the number of Bohmian trajectories and z_j and Z_A are weighting factors to preserve the norm of the states for the electrons and the nuclei, respectively.

Substituting Eqs. (18.78) and (18.79) into Eqs. (18.74) and (18.75) and assigning a separate guiding wave to each Bohmian trajectory, one can transform the nonlinear Hartree equations (18.74) and (18.75) into a set of coupled linear Schrödinger equations for the guiding waves,

$$i\hbar\frac{\partial}{\partial t}\chi_A^\alpha(\mathbf{R}_A,t) = \left[\frac{-\hbar^2}{2M_A}\nabla_A^2 + \sum_{B\neq A} V_{n-n}^{eff}(\mathbf{R}_A - \mathbf{R}_B^\alpha(t)) + \sum_j V_{n-e}^{eff}(\mathbf{R}_A - \mathbf{r}_j^\alpha(t)) + V_{ext}(\mathbf{R_A},t)\right]\chi_A^\alpha(\mathbf{R}_A,t) \tag{18.80}$$

and

$$i\hbar\frac{\partial}{\partial t}\Phi_i^\alpha(\mathbf{r}_i,t) = \left[\frac{-\hbar^2}{2m}\nabla_i^2 + \sum_{j\neq i} V_{e-e}^{eff}(\mathbf{r}_i - \mathbf{r}_j^\alpha(t)) + \sum_B V_{e-n}^{eff}(\mathbf{r}_i - \mathbf{R}_B^\alpha(t)) + V_{ext}(\mathbf{r_i},t)\right]\Phi_i^\alpha(\mathbf{r}_i,t), \tag{18.81}$$

where the non-local effective potentials are calculated as sums over the smoothed trajectory distributions,

$$V_{n-n}^{eff}(\mathbf{R}_A - \mathbf{R}_B^\alpha(t)) = \frac{1}{Z_B^\alpha}\sum_{\beta=1}^{M} V_{n-n}(\mathbf{R}_A - \mathbf{R}_B^\beta(t)) \exp\left[-\frac{|\mathbf{R}_B^\beta(t) - \mathbf{R}_B^\alpha(t)|^2}{\Sigma_B^\alpha(\mathbf{R}_B^\alpha,t)^2}\right], \tag{18.82}$$

$$V_{n-e}^{eff}(\mathbf{R}_A - \mathbf{r}_j^\alpha(t)) = \frac{1}{z_j^\alpha}\sum_{\beta=1}^{M} V_{n-e}(\mathbf{R}_A - \mathbf{r}_j^\beta(t)) \exp\left[-\frac{|\mathbf{r}_j^\beta(t) - \mathbf{r}_j^\alpha(t)|^2}{\sigma_j^\alpha(\mathbf{r}_j^\alpha,t)^2}\right], \tag{18.83}$$

$$V_{e-e}^{eff}(\mathbf{r}_i - \mathbf{r}_j^{\alpha}(t)) = \frac{1}{z_j^{\alpha}} \sum_{\beta=1}^{M} V_{e-e}(\mathbf{r}_i - \mathbf{r}_j^{\beta}(t)) \exp\left[-\frac{|\mathbf{r}_j^{\beta}(t) - \mathbf{r}_j^{\alpha}(t)|^2}{\sigma_j^{\alpha}(\mathbf{r}_j^{\alpha}, t)^2} \right], \tag{18.84}$$

$$V_{e-n}^{eff}(\mathbf{r}_i - \mathbf{R}_B^{\alpha}(t)) = \frac{1}{Z_B^{\alpha}} \sum_{\beta=1}^{M} V_{e-n}(\mathbf{r}_i - \mathbf{R}_B^{\beta}(t)) \exp\left[-\frac{|\mathbf{R}_B^{\beta}(t) - \mathbf{R}_B^{\alpha}(t)|^2}{\Sigma_B^{\alpha}(\mathbf{R}_B^{\alpha}, t)^2} \right], \tag{18.85}$$

where

$$z_j^{\alpha} = \sum_{\beta=1}^{M} \exp\left[-\frac{|\mathbf{r}_j^{\beta}(t) - \mathbf{r}_j^{\alpha}(t)|^2}{\sigma_j^{\alpha}(\mathbf{r}_j^{\alpha}, t)^2} \right], \tag{18.86}$$

$$Z_B^{\alpha} = \sum_{\beta=1}^{M} \exp\left[-\frac{|\mathbf{R}_B^{\beta}(t) - \mathbf{R}_B^{\alpha}(t)|^2}{\Sigma_B^{\alpha}(\mathbf{R}_B^{\alpha}, t)^2} \right]. \tag{18.87}$$

are the weighting factors. In fact, the effective potentials in Eqs. (18.82) and (18.83) describe the weighted non-local Coulomb interaction experienced by a given trajectory from the νth nuclear ensemble from the trajectories that belong to the ξth nuclear ensemble and from those from the jth electronic ensemble. The width of the Gaussian kernel $\Sigma_B^{\alpha}(\mathbf{R}_B^{\alpha}, t)$ plays the role of the characteristic length of the non-local quantum correlations that depend on the nuclear density (the density of trajectories) in the quantum system. At space locations where the trajectory density is higher the nuclear correlation length $\Sigma_B^{\alpha}(\mathbf{R}_B^{\alpha}, t)$ in Eqs. (18.82) and (18.85) is smaller in order to compensate for the higher number of interpolating Gaussians at that location. In these regions there are more intense interactions between the αth trajectory from the νth nuclear ensemble and the trajectories that represent the rest of electrons and nuclei. Because of the symmetry between the equations for electrons and nuclei (Eqs. (18.80) and (18.81)), similar considerations hold for the electronic non-local correlation length $\sigma_j^{\alpha}(\mathbf{r}_j^{\alpha}, t)$. The non-local correlation lengths $\Sigma_B^{\alpha}(\mathbf{R}_B^{\alpha}, t)$ and $\sigma_j^{\alpha}(\mathbf{r}_j^{\alpha}, t)$ are not free parameters and can be estimated using simple formulas

$$\Sigma_B^{\alpha}(\mathbf{R}_B^{\alpha}, t) = \Sigma \sqrt{\frac{G_B}{P_B^{\alpha}(\mathbf{R}, t)}}, \tag{18.88}$$

$$\sigma_j^{\alpha}(\mathbf{r}_j^{\alpha}, t) = \sigma \sqrt{\frac{g_j}{\rho_j^{\alpha}(\mathbf{r}, t)}}, \tag{18.89}$$

where $P_B^{\alpha}(\mathbf{R}, t)$ and $\rho_j^{\alpha}(\mathbf{r}, t)$ are pilot density estimates of the trajectory distributions for the ξth nucleus and the jth electron, which can be obtained using kernel density estimation with constant bandwidths Σ and σ, and G_B and g_j are the geometric means of the values of $P_B^{\alpha}(\mathbf{R}, t)$ and $\rho_j^{\alpha}(\mathbf{r}, t)$ for $\alpha = 1, ..., M$, respectively.

Equations (18.80) and (18.81) represent the quantum–quantum version of TDQMC-MD where both electronic and nuclear degrees are treated by sets of coupled Schrödinger equations and by the corresponding guiding equations for the Bohmian trajectories, Eqs. (18.76) and (18.77). Different approximations to the quantum–quantum description can be derived from Eqs. (18.80) and (18.81) by, for example, disentangling the system replicas for the nuclear degrees of freedom by letting $\Sigma_B^{\alpha} \to 0$ (this limit is known as the ultracorrelated nuclei) and then taking the classical limit of the nuclear degrees of freedom by taking the $\hbar \to 0$ limit.

Unlike other many-body quantum methods TDQMC does not involve calculation of overlap, exchange, and correlation integrals, which significantly improve its scaling properties. It also uses explicit Coulomb potentials instead of parametrized exchange-correlation potentials. The calculation of quantum potentials, which has been a major bottleneck for all particle methods, is avoided

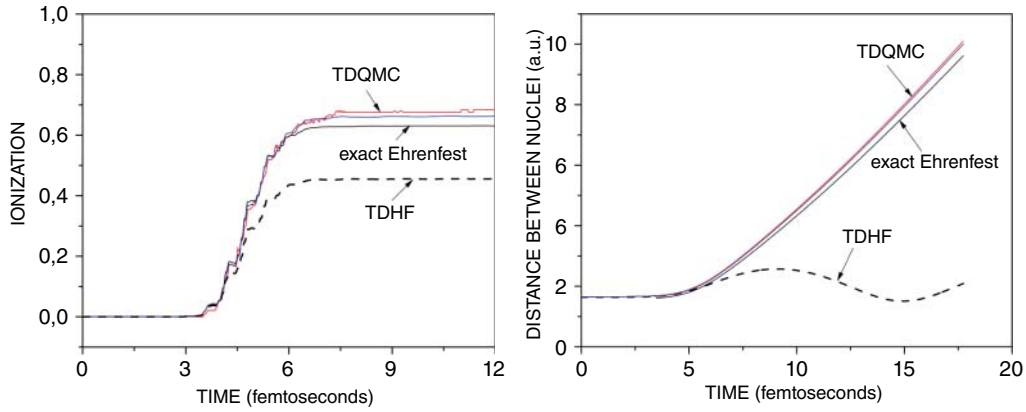


Figure 18.9 Left panel: Time-dependent ionization for one-dimensional hydrogen molecule in an external optical pulse with carrier frequency $\omega = 0.137$a.u. (335 nm) and peak intensity of $9 \cdot 10^{14}$W cm^{-2}. Black solid line: Ehrenfest MD; blue line: quantum-classical TDQMC result; red line: quantum-quantum TDQMC result; dashed line: TDHF. Reprinted from [I.P. Christov, *J. Chem. Phys.* 129, 214107 (2008)], with the permission of AIP.

in TDQMC. The model calculations for the dynamics of low-dimensional hydrogen molecules in external fields reveal that the TDQMC predictions for the ionization and for the internuclear distance are slightly enhanced as compared with the results from the exact Ehrenfest molecular dynamics (see Chapter 15), but they differ significantly from the TDHF results (see Figure 18.9). It is also remarkable that the TDQMC method is easily parallelized and requires little communication between the processors only for calculation of the non-local quantum correlation effects.

18.6 Conclusions

In this chapter we made a separation between approaches that are derived starting from the Born–Huang representation of the molecular wave function, and approaches that can be derived starting from the full (electron–nuclear) position-space representation of the time-dependent molecular Schrödinger equation. While in this work, we mainly focused our investigation on the use of Bohmian dynamics as a numerical approach for the reconstruction of the quantum dynamics of complex systems, it is worth mentioning that in the so-called analytic approach, Bohmian trajectories can also be reconstructed 'a posteriori' from the evolution of the system wave function becoming a new interpretative paradigm for the description of interesting quantum phenomena such as self-interference and entanglement [90].

In the Born–Huang basis electrons are described through a time-independent Schrödinger equation that is parametrically dependent on the nuclear degrees of freedom. Time-dependent wave packets representing the nuclear motion are discretized into a set of Bohmian fluid elements and are followed in time by integrating coupled equations of motion that are solved in the Lagrangian picture of fluid motion according to the Hamilton–Jacobi equations. Some of these techniques are already mature in the field of non-adiabatic molecular dynamics and have been applied to describe very large molecules, mainly thanks to their suitability to being coupled with well established electronic structure methods like multi-configurational self-consistent field (CASSCF) with or without corrections to include dynamical correlation [91], Coupled Cluster [92], and time-dependent density functional theory (TDDFT) [48, 93, 94]. Certainly, Bohmian

approaches based on the Born–Huang basis require the propagation of the exact Hamilton–Jacobi equation and hence cannot avoid the cumbersome computation of the quantum potential. Only under certain approximations can this problem be resolved [45, 46, 69, 86, 96]. Unfortunately, since the quantum potential does carry crucial information about the quantum nature of the nuclei, these approaches often fail to capture quantum nuclear effects such as tunneling, interference or the splitting of the nuclear probability density.

Alternatively, in Bohmian approaches based on the position-space representation of the full molecular wave function electrons and nuclei are treated in a symmetric way and are both represented by time-dependent equations of motion. Time-dependent electrons (nuclei) are described by waves that parametrically depend, via the Coulombic potential energy of the system, on nuclear (electronic) trajectories. These approaches allow to bypass the, typically necessary, computation of multiple BOPESs and NACTs and hence offer a new and attractive route to calculate observables and time correlation functions without relying on the widely used Born–Oppenheimer picture. Furthermore, the calculation of quantum potentials, which is a major bottleneck for methods based on the Born–Huang basis, is also avoided in these approaches. Despite the potential of these approaches to capture quantum nuclear effects, they have only been applied to model systems of very small molecules and their extension to systems made of more than a few atoms remains questionable mainly due to the lack of a proper procedure to couple them with well established electronic structure/dynamics methods [26, 72–74, 76, 77]. Overcoming these drawbacks may be just a matter of time, and would result in a prominent computational tool to describe non-adiabatic molecular dynamics.

References

1 The term *Copenhagen interpretation* refers to a set of rules for interpreting quantum phenomena devised by Born, Bohr, Heisenberg and others [98–100].

2 Feynman, R.P. and Hibbs, A.R. (1965). *Quantum Mechanics and Path Integrals*. New York: McGraw-Hill.

3 We chose the name *Bohmian mechanics* when referring to the work of Louis de Broglie and David Bohm because it is perhaps one of the most widespread names nowadays [101–104].

4 de Broglie, L. (1925). Recherches sur la theorie des quantas. *Ann. Phys.* 3: 22.

5 de Broglie, L. (1927). La mecanique ondulatorie et la structure atomique de la matiere et du ray- onnement. *J. Phys. Radium* 8: 225.

6 Bohm, D. (1952). A suggested interpretation of the quantum theory in terms of "hidden" variables. *Phys. Rev.* 85: 166–179.

7 Bohm, D. (1952). A suggested interpretation of the quantum theory in terms of "hidden" variables. *Phys. Rev.* 85: 180.

8 Bohm, D. (1953). Proof that probability density approaches $|\psi|^2$ in causal interpretation of quantum theory. *Phys. Rev.* 89: 458.

9 Ruggenthaler, M., Tancogne-Dejean, N., Flick, J. et al. (2018). From a quantum-electrodynamical light-matter description to novel spectroscopies. *Nat. Rev. Chem.* 2: 0118.

10 Tully, J.C. (2012). Perspective: nonadiabatic dynamics theory. *J. Chem. Phys.* 137: 22.

11 Hammes-Schiffer, S. and Soudackov, A.V. (2008). Proton-coupled electron transfer in solution, proteins, and electrochemistry. *J. Phys. Chem. B* 112: 14108–14123. PMID: 18842015.

12 Aviram, A. and Ratner, M.A. (1974). Molecular rectifiers. *Chem. Phys. Lett.* 29: 277–283.

13 McLachlan, A.D. (1964). A variational solution of the time-dependent Schrodinger equation. *Mol. Phys.* 8: 39–44.

14 Micha, D.A. (1983). A selfconsistent eikonal treatment of electronic transitions in molecular collisions. *J. Chem. Phys.* 78: 7138–7145.

15 Kirson, Z., Gerber, R.B., Nitzan, A., and Ratner, M.A. (1984). Dynamics of metal electron excitation in atom-surface collisions: a quantum wave packet approach. *Surf. Sci.* 137: 527–550.

16 Sawada, S.-I., Nitzan, A., and Metiu, H. (1985). Mean-trajectory approximation for charge- and energy-transfer processes at surfaces. *Phys. Rev. B* 32: 851–867.

17 Tully, J.C. and Preston, R.K. (1971). Trajectory surface hopping approach to nonadiabatic molecular collisions: the reaction of H^+ with D_2. *J. Chem. Phys.* 55: 562–572.

18 Tully, J.C. (1990). Molecular dynamics with electronic transitions. *J. Chem. Phys.* 93: 1061–1071.

19 Arasaki, Y., Takatsuka, K., Wang, K., and McKoy, V. (2003). Pump-probe photoionization study of the passage and bifurcation of a quantum wave packet across an avoided crossing. *Phys. Rev. Lett.* 90: 248303.

20 Landry, B.R. and Subotnik, J.E. (2012). How to recover marcus theory with fewest switches surface hopping: add just a touch of decoherence. *J. Chem. Phys.* 137: 22A513.

21 Horenko, I., Salzmann, C., Schmidt, B., and Schiitte, C. (2002). Quantum classical liouville approach to molecular dynamics: surface hopping gaussian phase-space packets. *J. Chem. Phys.* 117: 11075–11088.

22 Beck, M.H., Jckle, A., Worth, G.A., and Meyer, H.-D. (2000). The multiconfiguration time-dependent hartree (MCTDH) method: a highly efficient algorithm for propagating wave packets. *Phys. Rep.* 324: 1–105.

23 Domcke, W. and Stock, G. (2007). Theory of ultrafast nonadiabatic excited-state processes and their spectroscopic detection in real time. In: *Adv. Chem. Phys.*, 1–169. Wiley.

24 Lopreore, C.L. and Wyatt, R.E. (1999). Quantum wave packet dynamics with trajectories. *Phys. Rev. Lett.* 82: 5190–5193.

25 Born, M. and Oppenheimer, R. (1927). Zur quantentheorie der molekeln. *Ann. Phys.* 389: 457–484.

26 Albareda, G., Abedi, A., Tavernelli, I., and Rubio, A. (2016). Universal steps in quantum dynamics with time-dependent potential-energy surfaces: beyond the born-oppenheimer picture. *Phys. Rev. A* 94: 062511.

27 Benseny, A., Albareda, G., Sanz, A.S. et al. (2014). Applied bohmian mechanics. *Eur. Phys. J. D* 68: 286.

28 Wyatt, R.E. (2005) Chap. Derivative propagation along quantum trajectories). *Quantum Dynamics with Trajectories*, 235–253. New York: Springer.

29 Sanz, A.S. and Miret-Artes, S. (2012). *A Trajectory Description of Quantum Processes. I. Fundamentals*, Lecture Notes in Physics, vol. 850. Berlin: Springer.

30 Sanz, A.S. and Miret–Artes, S. (2008). Interplay of causticity and vorticality within the complex quantum Hamilton–Jacobi formalism. *Chem. Phys. Lett.* 458: 239–243.

31 Goldfarb, Y., Degani, I., and Tannor, D.J. (2006). Bohmian mechanics with complex action: a new trajectory-based formulation of quantum mechanics. *J. Chem. Phys.* 125: 231103.

32 Sanz, A.S. and Miret-Artes, S. (2007). Comment on 'Bohmian mechanics with complex action: a new trajectory-based formulation of quantum mechanics' [*j. chem. phys.* **125**, 231103 (2006)]. *J. Chem. Phys.* 127: 197101.

33 Goldfarb, Y., Degani, I., and Tannor, D.J. (2007). Response to "Comment on 'Bohmian mechanics with complex action: a new trajectory-based formulation of quantum mechanics' " [*j. chem. phys.* **127**, 197101 (2007)]. *J. Chem. Phys.* 127: 197102.

34 Goldfarb, Y. and Tannor, D.J. (2007). Interference in bohmian mechanics with complex action. *J. Chem. Phys.* 127: 161101.

35 Goldfarb, Y., Schiff, J., and Tannor, D.J. (2007). Unified derivation of Bohmian methods and the incorporation of interference effects. *J. Chem. Phys.* 111: 10416–10421.

36 Goldfarb, Y., Degani, I., and Tannor, D.J. (2007). Semiclassical approximation with zero velocity trajectories. *Chem. Phys.* 338: 106–112.

37 Zamstein, N. and Tannor, D.J. (2014). Communication: overcoming the root search problem in complex quantum trajectory calculations. *J. Chem. Phys.* 140: 041105.

38 Tannor, D.J. (2006). *Introduction to Quantum Mechanics: A Time-Dependent Perspective*. Sausalito, CA: University Science Books.

39 Zamstein, N. and Tannor, D.J. (2012). Non-adiabatic molecular dynamics with complex quantum trajectories. I. The diabatic representation. *J. Chem. Phys.* 137: 22A517.

40 Zamstein, N. and Tannor, D.J. (2012). Non-adiabatic molecular dynamics with complex quantum trajectories. II. The adiabatic representation. *J. Chem. Phys.* 137: 22A518.

41 Trahan, C.J., Hughes, K., and Wyatt, R.E. (2003). A new method for wave packet dynamics: derivative propagation along quantum trajectories. *J. Chem. Phys.* 118: 9911.

42 Wyatt, R.E. (2005). *Quantum Dynamics with Trajectories*. New York: Springer.

43 Tully, J.C. (1998). Mixed quantum classical dynamics. *Faraday Discuss.* 110: 407.

44 Marx, D. and Hutter, J. (2009). *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*. Cambridge University Press.

45 Curchod, B.F.E., Tavernelli, I., and Rothlisberger, U. (2011). Trajectory-based solution of the nonadiabatic quantum dynamics equations: an on-the-fly approach for molecular dynamics simulations. *Phys. Chem. Chem. Phys.* 13: 3231–3236.

46 Curchod, B.F.E. and Tavernelli, I. (2013). On trajectory-based nonadiabatic dynamics: Bohmian dynamics versus trajectory surface hopping. *J. Chem. Phys.* 138: 184112.

47 Curchod, B.F.E., Rothlisberger, U., and Tavernelli, I. (2012). Excited state dynamics with quantum trajectories. *Chimia* 66: 174–177.

48 Curchod, B.F.E., Rothlisberger, U., and Tavernelli, I. (2013). Trajectory-based nonadiabatic dynamics with time-dependent density functional theory. *Chem. Phys. Chem.* 14: 1314–1340.

49 Born, M. (1951). *Gott. Nachr. math. phys Kl* 1.

50 Born, M. and Huang, K. (1954). *Dynamical Theory of Crystal Lattices*. Clarendon, Oxford.

51 Handy, N.C., Yamaguchi, Y., and Schaefer, H.F. III, (1986). The diagonal correction to the Born–Oppenheimer approximation: its effect on the singlet-triplet splitting of CH_2 and other molecular effects. *J. Chem. Phys.* 84: 4481.

52 Cencek, W., Rychlewski, J., Jaquet, R., and Kutzelnigg, W. (1998). Sub-microhartree accuracy potential energy surface for H_3^+ including adiabatic and relativistic effects. I. Calculation of the potential points. *J. Chem. Phys.* 108: 2831.

53 Bunker, P.R. (1968). The electronic isotope shift in diatomic molecules and the partial breakdown of the Born–Oppenheimer approximation. *J. Mol. Spectrosc.* 28: 422–443.

54 Jensen, J.O. and Yarkony, D.R. (1988). On the evaluation of non-Born–Oppenheimer interactions for Born–Oppenheimer wave functions. V. A body fixed frame approach. Applications to isotope effects on equilibrium geometries and the adiabatic correction for the X $^1\Sigma^+$ *a* state of LiH. *J. Chem. Phys.* 89: 975.

55 Polyansky, O.L. and Tennyson, J. (1999). Ab initio calculation of the rotation-vibration energy levels of H_3^+ and its isotopomers to spectroscopic accuracy. *J. Chem. Phys.* 110: 5056.

56 Note that *Born-Oppenheimer approximation* is often used for what here is called *Born–Oppenheimer adiabatic approximation* (see also Refs. [105–107]).

57 The classical limit proposed in this derivation is sometimes called the "canonical condition" for enforcing classical behavior. It is mainly a *mathematical* procedure with limited physical content. Alternative formulations with their physical implications can be found in different references [108–110]. In Bohmian dynamics for example, the classical limit is more properly defined by $\left[-\frac{\hbar^2}{2m}\sum_A \frac{1}{M_A}\frac{\nabla_A^2 \mathscr{A}_\zeta}{\mathscr{A}_\zeta}\right] \to 0$ [108].

58 The continuity equation can be easily obtained by multiplying on the left both sides of Eq. (29) by $2\mathscr{A}_\zeta$ and using the relation $\frac{\nabla_A S_\zeta}{M_A} = \boldsymbol{v}_j^A$. We therefore get the equation $\frac{\partial \rho_\zeta}{\partial t} = -\sum_A \frac{1}{M_A}\nabla_A \cdot (\mathscr{A}_\zeta^2 \nabla_A S_\zeta) = -\sum_A \nabla_A \cdot (\rho_\zeta \boldsymbol{v}_j^A) = -\sum_A \nabla_A \cdot \boldsymbol{j}_j^A$ where $\rho_\zeta = \mathscr{A}_\zeta^2$ and $\boldsymbol{j}_j^A$ is the nuclear probability flux for A.

59 Wyatt, R.E. (2005). *Quantum dynamics with trajectories: introduction to quantum hydrodynamics*. In: *Interdisciplinary Applied Mathematics*. Springer.

60 Curchod, B.F.E., Tavernelli, I., and Rothlisberger, U. (2011). Trajectory-based solution of the nonadiabatic quantum dynamics equations: an on-the-fly approach for molecular dynamics simulations. *Phys. Chem. Chem. Phys.* 13: 3231–3236.

61 Tavernelli, I. (2013). Ab initio-driven trajectory-based nuclear quantum dynamics in phase space. *Phys. Rev. A* 87: 042501.

62 Klimontovich, Y.L. (1967). *The Statistical Theory of Non-Equilibrium Processes in a Plasma*. MIT press.

63 Ichimaru, S. (1992). *Statistical Plasma Physics*. Addison-Wesley.

64 Littlejohn, R.G. (1986). *Rev. Phys.* 138: 193.

65 Perdew, J.P., Burke, K., and Ernzerhof, M. (1996). Generalized gradient approximation made simple. *Phys. Rev. Lett.* 77: 3865–3868.

66 Garashchuk, S. and Rassolov, V.A. (2008). Stable long-time semiclassical description of zero-point energy in high-dimensional molecular systems. *J. Chem. Phys.* 129: 024109.

67 Garashchuk, S. and Rassolov, V.A. (2003). Semiclassical dynamics with quantum trajectories: formulation and comparison with the semiclassical initial value representation propagator. *J. Chem. Phys.* 118: 2482–2490.

68 Rassolov, V.A. and Garashchuk, S. (2004). Bohmian dynamics on subspaces using linearized quantum force. *J. Chem. Phys.* 120: 6815–6825.

69 Garashchuk, S., Rassolov, V.A., and Schatz, G.C. (2006). Semiclassical nonadiabatic dynamics based on quantum trajectories for the O (^{3}P, ^{1}D) + H_2 system. *J. Chem. Phys.* 124: 244307.

70 Garashchuk, S. and Rassolov, V.A. (2007). Semiclassical nonadiabatic dynamics of NaFH with quantum trajectories. *Chem. Phys. Lett.* 446: 395–400.

71 Garashchuk, S. and Rassolov, V. (2007). Stabilization of quantum energy flows within the approximate quantum trajectory approach. *Chem. A Eur. J.* 111: 10251–10255.

72 Albareda, G., Appel, H., Franco, I. et al. (2014). Correlated electron-nuclear dynamics with conditional wave functions. *Phys. Rev. Lett.* 113: 083003.

73 Albareda, G., Bofill, J.M., Tavernelli, I. et al. (2015). Conditional Born–Oppenheimer dynamics: quantum dynamics simulations for the model porphine. *J. Phys. Chem. Lett.* 6: 1529–1535.

74 Albareda, G., Kelly, A., and Rubio, A. (2019). Nonadiabatic quantum dynamics without potential energy surfaces. *Phys. Rev. Materials*. 3 (2): 023803.

75 Christov, I.P. (2006). Correlated non-perturbative electron dynamics with quantum trajectories. *Opt. Express* 14: 6906–6911.

76 Christov, I.P. (2007). Time-dependent quantum Monte Carlo: preparation of the ground state. *New J. Phys.* 9: 70–70.

77 Christov, I.P. (2007). Time-dependent quantum Monte Carlo and the stochastic quantization. *J. Chem. Phys.* 127: 134110.

78 Christov, I.P. (2009). Polynomial-time-scaling quantum dynamics with time-dependent quantum Monte Carlo. *J. Phys. Chem. A* 113: 6016–6021.

79 Christov, I.P. (2010). Time dependent quantum Monte Carlo: principles and perspectives. In: *AIP Conference Proceedings*, vol. 1228, 379–392. AIP Publishing.

80 Christov, I.P. (2011). Correlated electron dynamics with time-dependent quantum Monte Carlo: three-dimensional helium. *J. Chem. Phys.* 135: 044120–044120–7.

81 Gindensperger, E., Meier, C., and Beswick, J.A. (2000). Mixing quantum and classical dynamics using Bohmian trajectories. *J. Chem. Phys.* 113: 9369–9372.

82 Gindensperger, E., Meier, C., and Beswick, J.A. (2000). Communications: mixing quantum and classical dynamics using Bohmian trajectories. *J. Chem. Phys.* 113: 9369–9372.

83 Meier, C. (2004). Mixed quantum-classical treatment of vibrational decoherence. *Phys. Rev. Lett.* 93: 173003.

84 Cruz-Rodrguez, L., Uranga-Pia, L., Martnez-Mesa, A., and Meier, C. (2018). Quantum dynamics modeled by interacting trajectories. *Chem. Phys.* 503: 39–49.

85 Garashchuk, S., Rassolov, V., and Prezhdo, O. (2010). Semiclassical Bohmian dynamics. In: *Reviews in Computational Chemistry*, vol. 27 (ed. K.B. Lipkowitz), 287–368. New York: Wiley.

86 Prezhdo, O.V. and Brooksby, C. (2001). Quantum backreaction through the Bohmian particle. *Phys. Rev. Lett.* 86: 3215–3219.

87 Elsayed, T.A., Mølmer, K., and Bojer Madsen, L. (2017). Solving the Quantum Many-Body Problem with Bohmian Trajectories. ArXiv e-prints, arXiv:1706.00818 [quant-ph].

88 Shin, S. and Metiu, H. (1995). Nonadiabatic effects on the charge transfer rate constant: a numerical study of a simple model system. *J. Chem. Phys.* 102: 9285–9295.

89 Christov, I.P. (2008). Molecular dynamics with time dependent quantum Monte Carlo. *J. Chem. Phys.* 129: 214107.

90 Tavernelli, I. (2018). On the self-interference in electron scattering: Copenhagen, Bohmian and geometrical interpretations of quantum mechanics. *Ann. Phys.* 393: 447–465.

91 Jensen, F. (2007). *Introduction to Computational Chemistry*. Wiley.

92 Christiansen, O., Koch, H., and Jørgensen, P. (1995). The second-order approximate coupled cluster singles and doubles model CC2. *Chem. Phys. Lett.* 243: 409–418.

93 Runge, E. and Gross, E.K.U. (1984). Density-functional theory for time-dependent systems. *Phys. Rev. Lett.* 52: 997–1000.

94 Casida, M.E. (1995). Time-dependent density-functional response theory for molecules. In: *Recent Advances in Density Functional Methods* (ed. D.P. Chong), 155. Singapore: World Scientific.

95 Tavernelli, I. (2015). Nonadiabatic molecular dynamics simulations: synergies between theory and experiments. *Acc. Chem. Res.* 48: 792–800.

96 Gu, B. and Garashchuk, S. (2016). Quantum dynamics with gaussian bases defined by the quantum trajectories. *Chem. A Eur. J.* 120: 3023–3031.

97 Born, M. (1926). Zur quantenmechanik der stovorgange. *Z. Phys.* 37: 863.

98 Born, M., Heisenberg, W., and Jordan, P. (1925). Zur quantenmechanik ii. *Z. Phys.* 35: 557. english translation in Ref. [111].

99 Born, M. and Jordan, P. (1925). Zur quantenmechanik. *Z. Phys.* 34: 858. English translation in Ref. [111].

100 Cushing, J.T., Fine, A., and Goldstein, S. (eds.) (1996). *Bohmian Mechanics and Quantum Theory: An Appraisal*. Dordrecht: Kluwer Academic Publishers.

101 Oriols, X. and Mompart, J. (eds.) (2011). *Applied Bohmian Mechanics: From Nanoscale Systems to Cosmology*. Singapore: Pan Stanford Publishing.

102 Durr, D. and Teufel, S. (2009). *Bohmian Mechanics: The Physics and Mathematics of Quantum Theory*. Germany: Springer.

103 Durr, D., Goldstein, S., and Zanghi, N. (2012). *Quantum Physics without Quantum Philosophy*. Germany: Springer.

104 Ballhausen, C.J. and Hansen, A.E. (1972). Electronic spectra. *Annu. Rev. Phys. Chem.* 23: 15–38.

105 Worth, G.A. and Cederbaum, L.S. (2004). Beyond Born–Oppenheimer: molecular dynamics through a conical intersection. *Annu. Rev. Phys. Chem.* 55: 127–158.

106 Domcke, W., Yarkony, D., and Koppel, H. (eds.) (2004). *Conical Intersections: Electronic Structure, Dynamics & Spectroscopy*, vol. 15. World Scientific Pub Co Inc.

107 Bolivar, A.O. (2004). *Quantum-Classical Correspondence, Dynamical Quantization and the Classical Limit*. Berlin Heidelberg: Springer-Verlag.

108 Bowman, G.E. (2005). On the classical limit in Bohms theory. *Found. Phys.* 35: 605–625.

109 Holland, P.R. (1993). *The Quantum Theory of Motion – An Account of the de Broglie-Bohm Causal Interpretation of Quantum Mechanics*. Cambridge University Press.

110 van der Waerden, B.L. (1968). *Sources of Quantum Mechanics*. New York: Dover Publications.

19

Semiclassical Molecular Dynamics for Spectroscopic Calculations

Riccardo Conte and Michele Ceotto

Dipartimento di Chimica, Università degli Studi di Milano, via Golgi 19, 20133 Milano, Italy

Abstract

We present some historical and recently developed techniques to perform semiclassical spectroscopy calculations with both ground and excited state dynamics. The illustrated topics begin with a derivation of the basic semiclassical van Vleck propagator starting from Feynman's path integral formulation, followed by the description of the initial value representation formalism and a derivation of the Heller–Herman–Kluk–Kay semiclassical propagator. The chapter continues by introducing the time averaging technique and its very recent developments consisting in the multiple coherent, divide-and-conquer, and mixed semiclassical approaches. The main features of each method are described through examples with the intent of helping readers have a gentle learning curve. The chapter ends with a workflow chart, a few representative applications, a summary, and some conclusions.

19.1 Introduction

The primary research goal in the field of spectroscopy is the study of the interaction between light and matter. Depending on the frequency of the radiation, the term spectroscopy is usually specified as ultraviolet (UV), infrared (IR), microwave, and others. The different energy ranges of the radiation actually translate into different targets of investigation. For instance, UV spectroscopy focuses on electronic transitions, IR or Raman spectroscopies on molecular vibrations, and microwave spectroscopy on molecular rotations.

Infrared spectroscopy experiments are important tools for chemical characterization of unknown samples. In fact, peak positions in an IR spectrum are correlated to the molecular structures of the unknown chemical species. Many molecules strongly absorb in the mid-infrared region, defined by frequencies in the range between approximately 400 and 4000 cm^{-1}, with characteristic spectral patterns that can be compared to those of the unknown sample. The spectrum can also be employed in a quantitative way by exploiting peak intensities to estimate the species concentrations once a reference has been set. Furthermore, the width of the peaks can provide useful insights about the environment in which the molecule is embedded including pH conditions and the presence of hydrogen bonding. Finally, the experimental technique is fast and has enough sensitivity to require just a minimum amount of material.

Quantum Chemistry and Dynamics of Excited States: Methods and Applications,
First Edition. Edited by Leticia González and Roland Lindh.

However, the assignment of experimental spectra may become difficult when the complexity or dimensionality of molecular systems increases. Theoretical simulations can provide the necessary support by identifying spectral features with more confidence and relating them to the actual molecular motion. To this end, every theoretical method has to reconcile two different objectives. First, it must describe quantum effects like zero-point energies, overtones, and resonances, which cannot be neglected in spectroscopy. Secondly, it should be sufficiently manageable to be employed also for high dimensional systems.

The starting point of quantum theoretical spectroscopy is the time-independent eigenvalue equation $\hat{H}\Psi(\vec{r},\vec{R}) = E\Psi(\vec{r},\vec{R})$ because, once eigenvalues and eigenvectors of the molecular Hamiltonian are known, frequencies of vibration and absorption spectra can be determined. The general molecular Hamiltonian (neglecting spin-orbit interactions, and indicating electronic coordinates with $\vec{r}$ and nuclear ones with $\vec{R}$) can be written as a sum of five terms

$$\hat{H} = \hat{T}_N(\vec{R}) + \hat{T}_e(\vec{r}) + \hat{V}_{eN}(\vec{r},\vec{R}) + \hat{V}_{NN}(\vec{R}) + \hat{V}_{ee}(\vec{r}) \tag{19.1}$$

corresponding to the kinetic energies of nuclei and electrons and to the electron–nuclei, nuclei–nuclei, and electron–electron interaction potential energies [1]. The Hamiltonian is not separable into a nuclear and an electronic part due to the presence of $\hat{V}_{eN}(\vec{r},\vec{R})$ which makes the original eigenvalue problem difficult to solve. This interaction term is large and cannot be neglected, so in order to separate nuclear and electronic motions an approximation must be invoked. This is known as the Born–Oppenheimer approximation (for details see Chapter 1). Solution of the electronic problem leads to the construction of the potential energy surface for the system. Since the problem is solved at a discrete number of nuclear configurations, an analytical form of the global surface is obtained only upon application of a fitting procedure, which may be quite a difficult task and often constitutes a research topic on its own [2]. Furthermore, for each electronic state a different potential energy surface can be constructed.

The eigenvalues of the vibrational Hamiltonian associated with a specific electronic state are the vibrational energy levels E_k, starting with the zero-point one (E_0), from which it is straightforward to compute the frequencies of all spectral transitions. For instance, the frequency of the transition between the ground state and the generic k state with energy E_k is given by the difference $(E_k - E_0)/\hbar$. As for intensities, if the eigenfunctions are available, then (at least in principle) all dipole matrix elements $\mu_{0k} = \langle\Psi_0|\hat{\mu}|\Psi_k\rangle$ could be calculated and, eventually, the absorption formula

$$I_{abs}(E) \propto E \sum_k \Delta(E - E_k + E_0)\ |\mu_{0k}|^2 \tag{19.2}$$

can be evaluated, where $\Delta(E - E_k + E_0)$ is a bell-shaped function (rigorously a Dirac δ) peaked at $E_k - E_0$. However, the calculation may become prohibitive when the density of vibrational states is very large, which, for high dimensional systems, may happen at low energies.

The goal of this second part of the book is to deal with dynamical approaches. Calculations of vibrational frequencies and absorption spectra can indeed be undertaken by means of a dynamical approach. Specifically, the eigenenergies of the vibrational Hamiltonian can be computed from the Fourier transform of the survival amplitude $\langle\Xi|\Xi(t)\rangle$ of a generic reference state $|\Xi\rangle$. In fact,

$$
\begin{aligned}
I(E) &= \frac{1}{2\pi\hbar}\int_{-\infty}^{+\infty} dt\; e^{\mathrm{i}Et/\hbar}\langle\Xi|\Xi(t)\rangle \\
&= \frac{1}{2\pi\hbar}\int_{-\infty}^{+\infty} dt\; e^{\mathrm{i}Et/\hbar}\langle\Xi|e^{-\mathrm{i}\hat{H}t/\hbar}|\Xi\rangle \\
&= \frac{1}{2\pi\hbar}\int_{-\infty}^{+\infty} dt\; e^{\mathrm{i}Et/\hbar}\sum_k |\langle\Xi|E_k\rangle|^2 e^{-\mathrm{i}E_k t/\hbar} \\
&= \sum_k |\langle\Xi|E_k\rangle|^2\delta(E-E_k). \qquad (19.3)
\end{aligned}
$$

Peaks in the plot of $I(E)$ obtained from numerical implementation of the first relation in Eq. (19.3) are located at the vibrational eigenenergies of the system. Quantum vibrational frequencies are then found easily by scaling the eigenenergies with respect to the ground state energy (i.e., the vibrational zero point energy). A similar expression can be employed for absorption spectra. In fact, in this latter case, the relevant formula is [3]

$$
I_{abs}(E) \propto E\int_{-\infty}^{+\infty} dt\; e^{\mathrm{i}Et/\hbar}\;\langle\hat{\mu}(0)\hat{\mu}(t)\rangle, \qquad (19.4)
$$

where the average is over the density matrix of the system.

In the case of a photo-absorption involving two different electronic states (in the low temperature limit) [4]

$$
\langle\hat{\mu}(0)\hat{\mu}(t)\rangle = \langle\Xi_{gs}|\hat{\mu}e^{-\mathrm{i}\hat{H}_f t/\hbar}\hat{\mu}e^{\mathrm{i}\hat{H}_i t/\hbar}|\Xi_{gs}\rangle, \qquad (19.5)
$$

where $|\Xi_{gs}\rangle$ is the ground vibrational state of the lower electronic state, while $\hat{H}_i$ and $\hat{H}_f$ are the nuclear Hamiltonians in the lower and upper electronic surface respectively. This leads to the following working formula

$$
I_{abs}(E) \propto E\int_{-\infty}^{+\infty} dt\; e^{\mathrm{i}(E+E_0)t/\hbar}\;\langle\Theta(0)|\Theta(t)\rangle, \qquad (19.6)
$$

where $|\Theta\rangle$ is obtained by applying the electronic transition moment μ to the ground vibrational state $|\Xi_{gs}\rangle$ in the starting electronic state, i.e., $|\Theta\rangle = \hat{\mu}|\Xi_{gs}\rangle$. E_0 in Eq. (19.6) is the energy of $|\Xi_{gs}\rangle$. Within the Condon approximation, the electronic transition moment is taken as a constant. In this way, the initial wave packet for the excited state dynamics ($|\Theta(0)\rangle$) is prepared in a non-stationary state. By evolving it on the excited electronic state, the vibronic absorption spectrum is obtained similarly to power spectra and shows peaks at the energy E of the vibronic transitions. In photoemission calculations formulae are similar. The two surfaces are treated symmetrically, but there is a cubic dependence on the energy in front of the integration [3].

Many different theoretical approaches to spectroscopy have been developed to calculate $I(E)$ and $I_{abs}(E)$. In this chapter we focus on semiclassical (SC) molecular dynamics for spectroscopic calculations [5]. In this context "semiclassical" and "semiclassical dynamics" refer to a set of theories and time-dependent approaches based on an approximate quantum propagator (i.e., the semiclassical propagator) dependent on classical quantities. Therefore, the hallmark of SC dynamics is represented by the possibility to extract quantum features from classical molecular dynamics simulations [6]. This is permitted by the mathematical structure of the semiclassical propagator, which is based on classically evolved trajectories with remarkable ease of computational needs. Due to this property, semiclassical dynamics is a promising tool for high-dimensional applications. Other advantages include the possibility to work in Cartesian coordinates and to use the potential energy

obtained from the electronic problem "as is" and without any further approximation. In particular the potential can be provided in the form of an analytical surface or calculated ab initio on-the-fly for the whole dynamics. Finally, being based on classical dynamics, semiclassical methods may provide a more intuitive picture with respect to quantum approaches.

In this chapter devoted to SC dynamics we illustrate first the derivation of the basic semiclassical propagator as a stationary-phase approximation to Feynman's path integral formulation of quantum mechanics. Then the Heller–Herman–Kluk–Kay propagator is derived. It serves as the starting point for developing the multiple coherent and divide-and-conquer techniques, which are necessary to extend the applicability of semiclassical spectroscopy to high dimensional systems. These techniques are described theoretically and through examples that guide the reader in their application. We introduce also another family of semiclassical propagators known as "thawed Gaussian propagators" and a promising approach to condensed phase spectroscopy, before moving to applications concerning vibrational and vibronic spectroscopy. The methods presented can be applied to studies involving both ground-state and excited-state dynamics. Some general conclusions end the chapter.

19.2 From Feynman's Path Integral to van Vleck's Semiclassical Propagator

Among the many alternative derivations of the semiclassical propagator [7–9], perhaps the most intuitive one originates from Feynman's path integral formulation of the exact quantum propagator [10]. For more information on Feynman's propagator see Chapter 20.

We start from the observation that the differential Schrödinger equation

$$\mathrm{i}\hbar\frac{\partial|\Xi\rangle}{\partial t} = \hat{H}|\Xi\rangle \;\rightarrow |\Xi(t)\rangle = e^{-\mathrm{i}\hat{H}t/\hbar}|\Xi(0)\rangle \tag{19.7}$$

can be written in the path integral form by projecting the state $|\Xi\rangle$ onto the coordinate $\vec{q}$-space

$$\Xi(\vec{q}\,'(t)) \equiv \langle\vec{q}\,'|\Xi(t)\rangle = \langle\vec{q}\,'|e^{-\mathrm{i}\hat{H}t/\hbar}|\Xi(0)\rangle = \int_{-\infty}^{+\infty} d\vec{q}_0\langle\vec{q}\,'|e^{-\mathrm{i}\hat{H}t/\hbar}|\vec{q}_0\rangle\langle\vec{q}_0|\Xi(0)\rangle. \tag{19.8}$$

Insertion of the identity $\int_{-\infty}^{+\infty} d\vec{q}_0|\vec{q}_0\rangle\langle\vec{q}_0|$ into Eq. (19.8) introduces the idea of quantum propagation as the summation over all possible values in $\vec{q}$ − space of the probability amplitude $\langle\vec{q}\,'|e^{-\mathrm{i}\hat{H}t/\hbar}|\vec{q}_0\rangle$ matrix times the state vector $\langle\vec{q}_0|\Xi(0)\rangle$. The focus for solving the quantum propagation is now shifted to a suitable representation of the probability amplitude, instead of the wave function calculation as originally in Eq. (19.7). The first step consists in breaking N times the total time-evolution interval into infinitesimal $\Delta t = t/N$ time slices

$$\langle\vec{q}\,'|e^{-\mathrm{i}\hat{H}t/\hbar}|\vec{q}_0\rangle = \lim_{\Delta t\to 0}\langle\vec{q}\,'|(e^{-\mathrm{i}\hat{H}\Delta t/\hbar})^N|\vec{q}_0\rangle = \lim_{\Delta t\to 0}\langle\vec{q}\,'|(e^{-\mathrm{i}\hat{T}\Delta t/\hbar}e^{-\mathrm{i}\hat{V}\Delta t/\hbar})^N|\vec{q}_0\rangle, \tag{19.9}$$

where $\hat{T}$ is the kinetic energy operator, $\hat{V}$ the potential energy one, and the last equality is often referred to as the Suzuki–Trotter decomposition formula. The power of N can be interpreted as a product of N terms, so the next step consists in the insertion of $N-1$ quantum mechanical identities

$$\begin{aligned}\langle\vec{q}\,'|e^{-\mathrm{i}\hat{H}t/\hbar}|\vec{q}_0\rangle = \lim_{\Delta t\to 0}\int_{-\infty}^{+\infty} d\vec{q}_1\ldots\int_{-\infty}^{+\infty} d\vec{q}_{N-1}\langle\vec{q}\,'|(e^{-\mathrm{i}\hat{T}\Delta t/\hbar}e^{-\mathrm{i}\hat{V}\Delta t/\hbar})|\vec{q}_{N-1}\rangle\\ \times\langle\vec{q}_{N-1}|(e^{-\mathrm{i}\hat{T}\Delta t/\hbar}e^{-\mathrm{i}\hat{V}\Delta t/\hbar})|\vec{q}_{N-2}\rangle\cdots\langle\vec{q}_1|(e^{-\mathrm{i}\hat{T}\Delta t/\hbar}e^{-\mathrm{i}\hat{V}\Delta t/\hbar})|\vec{q}_0\rangle.\end{aligned} \tag{19.10}$$

In the time limit $\Delta t \to 0$ the potential can be approximated to a constant, the value being that of the potential at the mid-point between the two coordinates of each bracket in Eq. (19.10). The effect is twofold. On the one hand the exponential involving the potential can be evaluated straightforwardly. On the other hand what is left corresponds to a product of free-particle probability amplitudes which is analytically known. In fact, for a single probability amplitude

$$\langle \vec{q}_{N-1}|(e^{-\mathrm{i}\hat{T}\Delta t/\hbar}e^{-\mathrm{i}\hat{V}\Delta t/\hbar})|\vec{q}_{N-2}\rangle = e^{-\mathrm{i}\frac{V(\vec{q}_{N-1})+V(\vec{q}_{N-2})}{2}\Delta t/\hbar}\langle \vec{q}_{N-1}|e^{-\mathrm{i}\hat{T}\Delta t/\hbar}|\vec{q}_{N-2}\rangle. \tag{19.11}$$

Eventually, Feynman derived the following expression for the coordinate representation of the quantum propagator of a system of mass m (in the first sum $q_N = q'$)

$$\begin{aligned}\langle \vec{q}'|e^{-\mathrm{i}\hat{H}t/\hbar}|\vec{q}_0\rangle &= \lim_{N\to\infty}\left(\frac{m}{2\pi\mathrm{i}\hbar\Delta t}\right)^{N/2}\int_{-\infty}^{+\infty}d\vec{q}_1\ldots\int_{-\infty}^{+\infty}d\vec{q}_{N-1}\\ &\exp\left[\frac{\mathrm{i}m\Delta t}{2\hbar}\sum_{K=1}^{N}\left(\frac{q_K-q_{K-1}}{\Delta t}\right)^2\right]exp\left[-\frac{\mathrm{i}\Delta t}{\hbar}\left(\frac{V(\vec{q}_0)}{2}+\frac{V(\vec{q}')}{2}+\sum_{k=1}^{N-1}V(\vec{q}_K)\right)\right]\\ &= \int \mathcal{D}[\vec{q}(t)]e^{\mathrm{i}S_t(\vec{q}',\vec{q}_0)/\hbar}\end{aligned} \tag{19.12}$$

$S_t(\vec{q}', \vec{q}_0)$ is the action along the path going from $\vec{q}_0$ to $\vec{q}'$ in time t. It is defined as the time integral of the difference between kinetic and potential energies. $\int \mathcal{D}[\vec{q}(t)]$ is a special functional measure proportional to the product $\prod_{i=1}^{N-1}\int_{-\infty}^{+\infty}d\vec{q}_i$. In Figure 19.1 the integration of Eq. (19.12) is pictorially represented, with the time interval on the x-axis and the dashed vertical lines which are examples of time slices. Composition of all possible $\vec{q}_\alpha$ values at each time slice accounts for all possible path integrations.

Some of the possible paths are reported as continuous lines in Figure 19.1. Paths can be of any type, including polygonal chains (red curves). In general, they are not classical paths. Identification of all possible paths is, however, a formidable task which needs to be eased by approximating Eq. (19.12) in an appropriate way.

A stationary phase approximation [7] to the quantum propagator leads to what is called the semiclassical propagator. In detail, the stationary phase approximation to an oscillatory 1-dimensional integral can be written as

$$\int_{-\infty}^{+\infty} dx\, e^{\mathrm{i}f(x)} \approx \sum_{\{x_j|df(x_j)/dx=0\}}\sqrt{\frac{2\pi\mathrm{i}}{d^2f(x_j)/dx^2}}e^{\mathrm{i}f(x_j)}. \tag{19.13}$$

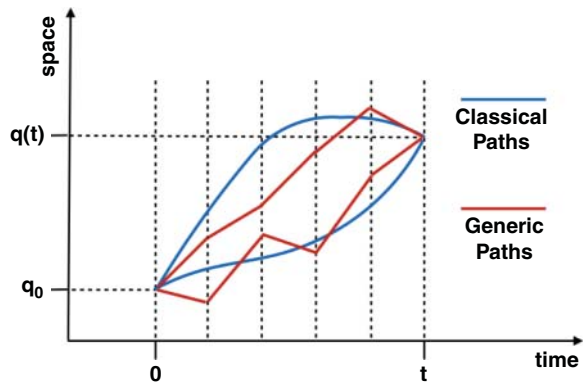


Figure 19.1 Pictorial representation of the Feynman path integral integration.

Easy extension to the multi-dimensional case is obtained by substituting $d^2f(x_j)/dx^2$ with the corresponding determinant of the matrix of second derivatives of f with respect to the position (i.e., the Hessian). This anticipates that in a SC simulation Hessian calculations are required along the trajectory. From classical mechanics it is known (Hamilton's principle) that for classical paths the functional derivative of the action with respect to a given path $\vec{q}(t)$ is zero. This is exactly the stationary phase condition for the path integral. Thus, the stationary phase approximation to the Feynman probability amplitude between $\vec{q}_0$ and $\vec{q}'$ in Eq. (19.12) leads to the following sum over all possible classical paths (called roots) connecting points $\vec{q}_0$ and $\vec{q}'$ in time t including fluctuations up to the second order around the classical action of each path

$$\langle \vec{q}' | e^{-\mathrm{i}\hat{H}t/\hbar} | \vec{q}_0 \rangle \propto \sum_{roots} e^{\mathrm{i}\left(S_t^{\mathrm{cl}}(\vec{q}',\vec{q}_0)+\frac{\partial}{\partial \vec{q}(t)}S_t^{\mathrm{cl}}(\vec{q}',\vec{q}_0)+\frac{\partial^2}{\partial \vec{q}(t)^2}S_t^{\mathrm{cl}}(\vec{q}',\vec{q}_0)\right)/\hbar}, \tag{19.14}$$

where $\partial S_t^{\mathrm{cl}}(\vec{q}',\vec{q}_0)/\partial\vec{q}(t) = 0$ because the sum is over classical paths. This is pictorially represented by the blue lines in Figure 19.1. Finding all the roots is not a trivial problem, which is generally known as the "double boundary problem". A suitable technique to overcome this formidable issue will be presented in the next section.

In the following we proceed with the algebra to determine explicitly an analytical formula for the coordinate representation of the semiclassical propagator. For simplicity we work in single dimensionality, but formulae are easily generalizable to multiple dimensions. The Feynman path integral is

$$\langle q' | e^{-\mathrm{i}\hat{H}t/\hbar} | q_0 \rangle = \lim_{N\to\infty} \left(\frac{Nm}{2\pi \mathrm{i}\hbar t}\right)^{N/2} \int_{-\infty}^{+\infty} dq_{N-1} \dots \int_{-\infty}^{+\infty} dq_1 e^{\mathrm{i}S_N(q)/\hbar}, \tag{19.15}$$

with $(q_N = q')$ and

$$S_N(q) = \frac{Nm}{2t}\sum_{K=1}^{N}(q_K - q_{K-1})^2 - \frac{t}{N}\left[\sum_{K=1}^{N-1} V(q_K) + \frac{1}{2}(V(q_0)+V(q_N))\right]. \tag{19.16}$$

The stationary phase condition is $\partial S_N/\partial q_K = 0$ for $K = 1, N-1$ leading to the following relation for each q_K

$$-V'(q_K) = m\frac{q_{K+1}+q_{K-1}-2q_K}{\Delta t^2} \approx m\ddot{q}_K, \tag{19.17}$$

where the second derivative of the position is approximated with a central finite difference formula. Equation (19.17) is nothing other than Newton's law for the classical motion of a particle of mass m moving in the potential V, thus confirming that application of the stationary phase condition restricts the general Feynman paths to classical trajectories only. To perform the stationary phase integration (see Eq. (19.13)), upon insertion of Eq. (19.16) into Eq. (19.15) one integrates the Gaussian integrals of each path fluctuation. By writing [11]

$$\left|\frac{\partial^2 S_N}{\partial q_i \partial q_j}\right|^{-1} = -\left(\frac{t}{Nm}\right)^N \left|\frac{\partial^2 S_N}{\partial q_0 \partial q_N}\right|, \tag{19.18}$$

the final, multidimensional expression is

$$\begin{aligned}\left\langle \vec{q}' \left| e^{-\frac{\mathrm{i}}{\hbar}\hat{H}t} \right| \vec{q}_0 \right\rangle &\approx \sum_{roots}\left[-\frac{\left|\frac{\partial^2 S_t}{\partial \vec{q}'\partial \vec{q}_0}\right|}{(2\pi\mathrm{i}\hbar)^F}\right]^{1/2} \frac{e^{\frac{\mathrm{i}}{\hbar}S_t(\vec{q}',\vec{q}_0)}}{e^{\mathrm{i}\upsilon\pi/2}} \\ &= \sum_{roots}\left[(2\pi\mathrm{i}\hbar)^F \left|\frac{\partial \vec{q}'}{\partial \vec{p}_0}\right|\right]^{-1/2} \frac{e^{\frac{\mathrm{i}}{\hbar}S_t(\vec{q}',\vec{q}_0)}}{e^{\mathrm{i}\upsilon\pi/2}}\end{aligned} \tag{19.19}$$

where we have introduced the formula valid for classical trajectories $\partial S_t(\vec{q}', \vec{q}_0)/\partial \vec{q}_0 = -\vec{p}_0$ and F indicates the number of degrees of freedom [12, 13]. In Eq. (19.19) the sum is only over classical trajectories.

Equation (19.19) represents the semiclassical approximation to Feynman's path integral [7], an expression analogous to the one proposed by van Vleck many years before [13]. The exponential term $e^{-i\upsilon\pi/2}$ ensures the continuity of the complex square root of the pre-exponential factor and υ is called the Maslov index [14]. It is important to stress that the semiclassical approximation is not simply a sum over all possible classical paths of the exact Feynman path integral, but each path takes into account the second-order fluctuations around the classical path. In other words, the semiclassical approximation is able to reproduce quantum effects with high accuracy because it does not only include the interference effects generated by the sum of many classical trajectories weighted by the complex quantity $\exp[iS(\vec{q}', \vec{q}_0)/\hbar]$, but it also reproduces the quantum fluctuations $\partial^2 S_t/\partial \vec{q}' \partial \vec{q}_0$ up to the second order around each path.

19.3 The Semiclassical Initial Value Representation and the Heller–Herman–Kluk–Kay Formulation

Equation (19.19) has the intriguing feature of reproducing quantum effects starting from classical trajectories, but nevertheless its application has been quite limited. The reason for this limitation is twofold.

First, simulations are hindered by the presence of caustic points (defined by the relation $\partial \vec{q}'/\partial \vec{p}_0 = 0$) at which the determinant in the pre-exponential factor becomes singular. $\vec{q}'$ is a caustic (or focal) point for the classical trajectories started at $\vec{q}_0$ when the trajectories reach $\vec{q}'$ independently of their initial momentum. As an example, we calculate the caustic points of a 1-dimensional harmonic oscillator of unitary mass that starts at (p_0, q_0). After an evolution time T

$$q(T) = q_0 cos(\omega T) + \frac{p_0}{\omega} sin(\omega T). \tag{19.20}$$

The condition for caustic points $\partial q(T)/\partial p_0 = 0$ is in this case equivalent to $sin(\omega T)/\omega = 0$. This means that at times $T = n\pi/\omega$ the trajectory lands on the caustic points, which are located at $\pm q_0$, i.e., the inversion points of the harmonic oscillator if no momentum is given initially.

Secondly, a search for multidimensional trajectories satisfying the double boundary condition is requested.

Even though the first of the two issues can be removed analytically by means of "uniformization" approximations, which consist in switching to appropriate, caustic-free representations in proximity of the singularity, solving the double boundary problem is quite a cumbersome task in spite of the ever increasing availability of computational power.

To overcome both issues, William H. Miller incorporated an initial value representation (IVR) into the SC approximation [15]. In semiclassical IVR (SCIVR) the sum over all possible boundary-ended classical trajectories is replaced by an integration over initial momenta, such that the final position is equal to $\vec{q}'$

$$\sum_{\text{roots}} \rightarrow \int d\vec{p}_0 \left| \frac{\partial \vec{q}'}{\partial \vec{p}_0} \right| \delta(\vec{q}_t - \vec{q}'). \tag{19.21}$$

In Eq. (19.21) the determinant is the Jacobian for the change of variables from $\vec{q}'$ (the final position) to $\vec{p}_0$ (the starting momentum), while the delta function enforces the desired boundary condition.

This change of variables presents two advantages. First, the pre-exponential singularity is removed and the semiclassical approximation is uniform. Secondly, the root search is replaced by a phase space integration, which can be evaluated by means of Monte Carlo techniques with much less computational effort. Eventually, the SCIVR version of Eq. (19.19) is

$$\left\langle \vec{q}' \left| e^{-\frac{\mathrm{i}}{\hbar}\hat{H}t} \right| \vec{q}_0 \right\rangle \approx \int d\vec{p}_0 \delta(\vec{q}_t - \vec{q}') \left[\frac{|\partial \vec{q}'/\partial \vec{p}_0|}{(2\pi\mathrm{i}\hbar)^F} \right]^{1/2} \frac{e^{\frac{\mathrm{i}}{\hbar}S_t(\vec{p}_0,\vec{q}_0)}}{e^{\mathrm{i}\upsilon\pi/2}}. \tag{19.22}$$

For numerical calculations, the Dirac delta in Eq. (19.22) can be conveniently represented as an appropriately narrowed Gaussian function.

However, the SCIVR probability amplitude is not employed *per se*, but rather for calculating physical observables. We will use it for spectroscopic calculations using the time-dependent approach of Eq. (19.3). By inserting twice the $\vec{q}$-space identity and by means of the SCIVR probability amplitude of Eq. (19.22), the survival amplitude becomes

$$\begin{aligned} \left\langle \Xi \left| e^{-\frac{\mathrm{i}}{\hbar}\hat{H}t} \right| \Xi \right\rangle &= \int d\vec{q}' d\vec{q}_0 \langle \Xi | \vec{q}' \rangle \left\langle \vec{q}' | e^{-\frac{\mathrm{i}}{\hbar}\hat{H}t} | \vec{q}_0 \right\rangle \langle \vec{q}_0 | \Xi \rangle \\ &\approx \int d\vec{p}_0 \int d\vec{q}_0 \left[\frac{1}{(2\pi\mathrm{i}\hbar)^F} \left| \frac{\partial \vec{q}_t}{\partial \vec{p}_0} \right| \right]^{1/2} \frac{e^{\frac{\mathrm{i}}{\hbar}S_t(\vec{p}_0,\vec{q}_0)}}{e^{\mathrm{i}\upsilon\pi/2}} \Xi^*(\vec{q}_t)\Xi(\vec{q}_0). \end{aligned} \tag{19.23}$$

As anticipated, the great advantage of Eq. (19.23) is that now the survival probability can be evaluated via Monte Carlo integration. Calculations aimed at vibrational spectroscopy are commonly performed in normal mode coordinates. By diagonalizing the Hessian matrix at the equilibrium geometry the normal frequencies of vibrations are obtained, and the associated eigenvectors serve to define the transformation matrix from normal modes to Cartesian coordinates.

Further advances in SC dynamics were introduced by Heller [16], who inspired later work by Herman and Kluk [17], and eventually Kay [18–20]. They represented the semiclassical propagator in terms of coherent states. Coherent states ($|\vec{p},\vec{q}\rangle$) have a Gaussian representation in coordinate space

$$\langle \vec{x} | \vec{p},\vec{q} \rangle = \left(\frac{\det(\Gamma)}{\pi^F} \right)^{1/4} e^{-(\vec{x}-\vec{q})^{\mathrm{T}} \frac{\Gamma}{2} (\vec{x}-\vec{q}) + \mathrm{i}\vec{p}^{\,\mathrm{T}}(\vec{x}-\vec{q})/\hbar}, \tag{19.24}$$

where the Gaussian width is determined by the (usually diagonal) Γ width parameter matrix. It is possible either to reformulate the Feynman paths directly in terms of coherent states [21, 22], or to represent Eq. (19.23) on the basis of coherent states [17]. In both instances the following expression for the survival amplitude is derived

$$\begin{aligned} \left\langle \Xi \left| e^{-\frac{\mathrm{i}}{\hbar}\hat{H}t} \right| \Xi \right\rangle \approx \frac{1}{(2\pi\hbar)^F} \iint d\vec{q}_0 d\vec{p}_0 C_t(\vec{p}_0,\vec{q}_0) e^{\frac{\mathrm{i}}{\hbar}S_t(\vec{p}_0,\vec{q}_0)} \\ \times \langle \Xi | \vec{p}_t,\vec{q}_t \rangle \langle \vec{p}_0,\vec{q}_0 | \Xi \rangle, \end{aligned} \tag{19.25}$$

where, in its most general form [18],

$$C_t(\vec{p}_0,\vec{q}_0) = \sqrt{\left| \frac{1}{2} \left(\frac{\partial \vec{q}_t}{\partial \vec{q}_0} + \Gamma^{-1} \frac{\partial \vec{p}_t}{\partial \vec{p}_0} \Gamma - \mathrm{i}\hbar \frac{\partial \vec{q}_t}{\partial \vec{p}_0} \Gamma + \frac{\mathrm{i}\Gamma^{-1}}{\hbar} \frac{\partial \vec{p}_t}{\partial \vec{q}_0} \right) \right|}. \tag{19.26}$$

The semiclassical way to calculate the survival amplitude of Eq. (19.25) numerically is to perform the phase space integration by Monte Carlo methods upon sampling of the initial phase space coordinates $(\vec{p}_0,\vec{q}_0)$. The real part of the term $\langle \vec{p}_0,\vec{q}_0 | \Xi \rangle$ in Eq. (19.25) constitutes a natural weight for the Monte Carlo sampling. After the classical evolution is performed employing preferentially a

simplectic algorithm [23], at the generic time t, starting from the values of $\vec{p}_t$ and $\vec{q}_t$, the classical action $S_t(\vec{p}_0, \vec{q}_0)$ and the pre-exponential factor $C_t(\vec{p}_0, \vec{q}_0)$ are calculated. Eventually, the power spectrum is obtained by Fourier transforming Eq. (19.25)

$$I(E) = \frac{1}{2\pi\hbar}\int_{-\infty}^{+\infty} dt e^{iEt/\hbar}\frac{1}{(2\pi\hbar)^F}\iint d\vec{q}_0 d\vec{p}_0 C_t(\vec{p}_0,\vec{q}_0) \times e^{iS_t(\vec{p}_0,\vec{q}_0)/\hbar}\langle\Xi|\vec{p}_t,\vec{q}_t\rangle\langle\vec{p}_0,\vec{q}_0|\Xi\rangle. \tag{19.27}$$

19.4 A Derivation of the Heller–Herman–Kluk–Kay Propagator

Once the mathematical expression of the Heller–Herman–Kluk–Kay (HHKK) propagator has been introduced

$$(e^{-i\hat{H}t/\hbar})_{\mathrm{HHKK}} = \frac{1}{(2\pi\hbar)^F}\iint d\vec{q}_0 d\vec{p}_0 C_t(\vec{p}_0,\vec{q}_0) e^{iS_t(\vec{p}_0,\vec{q}_0)/\hbar}|\vec{p}_t,\vec{q}_t\rangle\langle\vec{p}_0,\vec{q}_0|, \tag{19.28}$$

following Miller's original derivation [24] we detail how Eq. (19.25) – and consequently Eq. (19.28) – can be derived from Eq. (19.23).

For simplicity we work in one dimension and begin by showing that an appropriate filter can be introduced into an oscillatory integral to speed up the convergence of the integration. This technique, known as Filinov filtering, is based on the insertion of a Gaussian identity into the target integral. Specifically

$$I = \int dx_0 e^{i\phi(x_0)} = \int dx_0 \int dx_G \sqrt{\frac{A}{\pi}} e^{-A(x_G-x_0)^2} e^{i\phi(x_0)}, \tag{19.29}$$

which can be approximated by expanding to the second order the function $\phi(x_0)$ around x_G and then integrating analytically in the x_0 variable

$$\begin{aligned} I &\approx \sqrt{\frac{A}{\pi}}\int dx_G \int dx_0 e^{-A(x_0-x_G)^2} e^{i[\phi(x_G)+\phi'(x_G)(x_0-x_G)+\frac{1}{2}\phi''(x_G)(x_0-x_G)^2]} \\ &= \int dx_G\, e^{i\phi(x_G)} \sqrt{\frac{A}{A-\frac{i}{2}\phi''(x_G)}}\, e^{-\phi'(x_G)^2/[4A-2i\phi''(x_G)]}. \end{aligned} \tag{19.30}$$

Equation (19.30) retains the original integrand with the addition of a damping factor that facilitates numerical convergence [25]. The A parameter can be chosen arbitrarily, but Makri and Miller suggested employing the value $A = [i\phi''(x_G) + c^{-1}]/2$, where c is a constant or a constant matrix in the multidimensional case. This choice is justified by the observation that the Gaussian identity is still approximately valid even if A is a function of the variable of integration [26].

The next step consists in applying this Filinov filter to calculate the propagation from a coherent state $|\Xi_i\rangle \equiv |p_i, q_i\rangle$ to a coherent state $|\Xi_f\rangle \equiv |p_f, q_f\rangle$ (both of Γ width) by means of van Vleck's propagator according to

$$\left\langle \Xi_f \left| e^{-\frac{i}{\hbar}\hat{H}t} \right| \Xi_i \right\rangle = \int dq_t dq_0\, (2\pi i\hbar\, \partial q_t/\partial p_0)^{-1/2}\, e^{iS_t(q_t,q_0)/\hbar}\, \langle p_f, q_f|q_t\rangle\, \langle q_0|p_i,q_i\rangle, \tag{19.31}$$

where the exponential with the Maslov index is left implicit. We work out the case of a 1-dimensional system (i.e., a bidimensional phase space) but results are generalizable to multiple dimensions. We assume also that the complex-valued pre-exponential factor in Eq. (19.31) is slowly varying with respect to the rest of the integrand, so that the filter applies only to the latter. Then, by

rearranging Eq. (19.31) in the form $I = \int dq_t\, dq_0\, A'\, e^{i\phi(q_t,q_0)}$ with A' including the pre-exponential factor and normalization constants, we need to evaluate the first and second derivatives of $\phi(q_t,q_0)$

$$\phi(q_t,q_0) = S_t(q_t,q_0) + p_i(q_0-q_i) - p_f(q_t-q_f) + i\frac{\Gamma}{2}(q_f-q_t)^2 + i\frac{\Gamma}{2}(q_i-q_0)^2 \tag{19.32}$$

$$\frac{\partial\phi(q_t,q_0)}{\partial q_0} = -p_0 + p_i + i\Gamma(q_0-q_i) \qquad \frac{\partial\phi(q_t,q_0)}{\partial q_t} = p_t - p_f + i\Gamma(q_t-q_f) \tag{19.33}$$

$$\begin{pmatrix} \frac{\partial^2\phi(q_t,q_0)}{\partial q_t^2} & \frac{\partial^2\phi(q_t,q_0)}{\partial q_0\partial q_t} \\ \frac{\partial^2\phi(q_t,q_0)}{\partial q_t\partial q_0} & \frac{\partial^2\phi(q_t,q_0)}{\partial q_0^2} \end{pmatrix} = \begin{pmatrix} \frac{\partial p_t}{\partial q_t} + i\Gamma & \frac{\partial p_t}{\partial q_0} \\ -\frac{\partial p_0}{\partial q_t} & -\frac{\partial p_0}{\partial q_0} + i\Gamma \end{pmatrix}, \tag{19.34}$$

and then choose a diagonal Filinov parameter matrix

$$c = \begin{pmatrix} c_0 & 0 \\ 0 & c_1 \end{pmatrix}. \tag{19.35}$$

In agreement with Eq. (19.30) the Filinov filtered version of the integral in Eq. (19.31) becomes

$$I \approx \int dq_t dq_0 \left(\frac{\Gamma}{\pi}\right)^{1/2} \left(2\pi i\hbar\frac{\partial q_t}{\partial p_0}\right)^{-1/2} det[1+ic\phi'']^{1/2}\, e^{[-\phi'^T\frac{c}{2}\phi'+i\phi]}, \tag{19.36}$$

where ϕ' is shorthand for the vector of first derivatives (Eq. (19.33)) and ϕ'' for the matrix of second derivatives in Eq. (19.34). Eventually, moving to the IVR framework

$$I \approx (2\pi\hbar)^{-1} \int dp_0 dq_0 \left(-2i\hbar\Gamma\frac{\partial q_t}{\partial p_0}\right)^{1/2} C_{\text{Fil}}\, e^{[i\phi+\phi_{\text{Fil}}]}. \tag{19.37}$$

Matrix–matrix and matrix–vector products appearing in Eq. (19.36) can be evaluated using standard algebraic manipulations. This leads to

$$C_{\text{Fil}}^2 = (1-c_0\Gamma)(1-c_1\Gamma) + \left(\frac{\partial q_t}{\partial p_0}\right)^{-1}\left(ic_1\frac{\partial q_t}{\partial q_0}(1-c_0\Gamma) + ic_0(1-c_1\Gamma)\frac{\partial p_t}{\partial p_0} - c_0c_1\frac{\partial p_t}{\partial q_0}\right)$$

$$\phi_{\text{Fil}} = -\frac{c_1}{2}[(p_t-p_f)^2 + 2i\Gamma(p_t-p_f)(q_t-q_f) - \Gamma^2(q_t-q_f)^2]$$
$$-\frac{c_0}{2}[(p_0-p_i)^2 + 2i\Gamma(p_i-p_0)(q_0-q_i) - \Gamma^2(q_0-q_i)^2],$$

and the HHKK propagator of Eq. (19.28) in the mono-dimensional case is obtained by choosing $c_0 = c_1 = 1/2\Gamma$. This demonstrates that the more practical HHKK propagator can be derived as an approximate version of the van Vleck one. The key advantages of the former lie in the built-in Gaussian probability density for Monte Carlo integration and in the possibility to preserve the unitarity of the propagator longer in time thanks to its pre-exponential factor [27].

19.5 The Time-Averaging Filter

Application of the original HHKK propagator is limited to low-dimensional systems due to the presence of the oscillatory exponential in the integrand. Several methods such as cellular dynamics [28–30], Filinov and generalized Filinov filtering [26, 31, 32] have been proposed and can be effective to speed up the convergence of the Monte Carlo integration of Eq. (19.27). Here we describe in

detail a different approach based on the insertion of a time-averaging (TA) filter. Time-averaging was introduced by Elran and Kay [33, 34], and later by Kaledin and Miller [35, 36]. The idea is that starting from a phase space average of the form of the HHKK propagator

$$I_{\text{PS}} = \iint d\vec{p}_0 d\vec{q}_0 A(\vec{p}_0, \vec{q}_0), \tag{19.38}$$

one can speed up convergence by time-averaging the integrand, thus damping disadvantageous oscillations. In practice, the time averaged version of I_{PS} is

$$I_{\text{PS-TA}} = \iint d\vec{p}_0 d\vec{q}_0 \frac{1}{T} \int_0^T dt\, A(\vec{p}_t, \vec{q}_t), \tag{19.39}$$

which is fully equivalent to I_{PS}. The demonstration is achieved first by changing the order of integrations, and then by invoking Liouville's theorem, which guarantees that the phase-space distribution function is constant along the trajectories and the change of variables $d\vec{p}_0 d\vec{q}_0 \rightarrow d\vec{p}_t d\vec{q}_t$ has a unitary Jacobian.

$$\begin{aligned} I_{\text{PS-TA}} &= \frac{1}{T} \int_0^T dt \iint d\vec{p}_0 d\vec{q}_0 A(\vec{p}_t, \vec{q}_t) \\ &= \frac{1}{T} \int_0^T dt \iint d\vec{p}_t d\vec{q}_t A(\vec{p}_t, \vec{q}_t) \\ &= \frac{1}{T} \int_0^T dt I_{PS} = I_{PS} \end{aligned} \tag{19.40}$$

The phase-space and time integrations commute, but Eq. (19.40) is exact only when, within numerical accuracy, both integrations have converged. Kaledin and Miller [35, 36] worked out the following time averaged version of Eq. (19.27)

$$\begin{aligned} I(E) = \frac{1}{(2\pi\hbar)^F} \iint d\vec{q}_0 d\vec{p}_0 \frac{1}{T} \int_0^T dt_1 \frac{Re}{\pi\hbar} \int_0^{+\infty} dt e^{iEt/\hbar} \\ \times C_{t_1+t}(\vec{p}_{t_1}, \vec{q}_{t_1}) e^{iS_{t_1+t}(\vec{p}_{t_1}, \vec{q}_{t_1})/\hbar} \langle \Xi | \vec{p}_{t_1+t}, \vec{q}_{t_1+t} \rangle \langle \vec{p}_{t_1}, \vec{q}_{t_1} | \Xi \rangle. \end{aligned} \tag{19.41}$$

Finally, moving from the integration variable t to $t_2 = t_1 + t$, Eq. (19.41) becomes

$$\begin{aligned} I(E) = \frac{1}{(2\pi\hbar)^F} \iint d\vec{q}_0 d\vec{p}_0 \frac{Re}{\pi\hbar T} \int_0^T dt_1 \\ \times \int_{t_1}^{+\infty} dt_2 e^{\frac{i}{\hbar}(S_{t_2}(\vec{p}_0, \vec{q}_0) + Et_2)} \langle \chi | \vec{p}_{t_2}, \vec{q}_{t_2} \rangle e^{-\frac{i}{\hbar}(S_{t_1}(\vec{p}_0, \vec{q}_0) + Et_1)} \\ \times \langle \vec{p}_{t_1}, \vec{q}_{t_1} | \chi \rangle C_{t_2}(\vec{p}_{t_1}, \vec{q}_{t_1}) \end{aligned} \tag{19.42}$$

which is the time-averaged SCIVR (TA-SCIVR) formula for power spectrum calculations.

We now focus on the bound states and consider a number N_{vib} of vibrational degrees of freedom. In order to get to a simpler form of Eq. (19.42), we approximate the pre-exponential factor $C_{t_2}(\vec{p}_{t_1}, \vec{q}_{t_1})$ in Eq. (19.42), in agreement with previous work [36]. We note that for the harmonic oscillator $C_{t_2}(\vec{p}_{t_1}, \vec{q}_{t_1}) = e^{-i(\hbar\omega)(t_2 - t_1)/2\hbar} = e^{-i(\phi(t_1) - \phi(t_2))/\hbar}$, which is a complex number of unit modulus with a time-dependent phase. The pre-exponential factor is analytically separable, i.e., $C_{t_2}(\vec{p}_{t_1}, \vec{q}_{t_1}) = C_{t_2}(\vec{p}_{t_0}, \vec{q}_{t_0}) C^*_{t_1}(\vec{p}_{t_0}, \vec{q}_{t_0})$, where $C_t(\vec{p}_0, \vec{q}_0) \approx e^{i\phi(t)/\hbar}$. Inspired by this consideration, we decide to approximate for any bound degree of freedom the exact Herman Kluk prefactor to an element of unitary norm dependent on the prefactor phase $C_t(\vec{p}_0, \vec{q}_0) \approx e^{i\phi(t)/\hbar}$, where $\phi(t) = \text{phase}[C_t(\vec{p}_0, \vec{q}_0)]$, as suggested in Ref. [36]. Finally, noting that the integration over t_2 is for

practical purposes limited to the simulation time T and duplicating the integral by integrating from 0 to T also in t_2 (and consequently dividing the result by a factor of 2), the power spectrum formula obtained by means of this separable approximation is

$$I(E) = \left(\frac{1}{2\pi\hbar}\right)^{N_{vib}} \iint d\vec{p}_0 d\vec{q}_0 \frac{1}{2\pi\hbar T} \times \left| \int_0^T dt e^{\frac{i}{\hbar}[S_t(\vec{p}_0,\vec{q}_0)+Et+\phi_t]} \langle \Xi | \vec{p}_t \vec{q}_t \rangle \right|^2 . \tag{19.43}$$

Equation (19.43) is the "separable" version of TA SCIVR. For the phase-space integration in Eq. (19.43) the quantity $|\langle \vec{p}_0, \vec{q}_0 | \Xi \rangle|^2$ is often employed as Monte Carlo importance sampling density and named "Husimi distribution", if the reference state $|\Xi\rangle$ is chosen to be a coherent state. The choice of this Gaussian sampling density is, however, arbitrary and it can be tuned by varying the Gaussian width. The convergence issue in TA SCIVR is much alleviated since a number of trajectories of the order of just 1000 per degree of freedom is generally required, and the approach has permitted evaluation of vibrational frequencies accurately for a set of small isolated molecules including water, formaldehyde, and methane [37]. For such molecules pure quantum dynamical calculations have been also performed, so the challenge for semiclassical dynamics is to investigate much larger systems. However, for this purpose it is necessary to develop the semiclassical theory further in order to obtain reliable results on the basis of just a handful of classical trajectories.

19.6 The Multiple Coherent States SCIVR

The time-averaged version of the semiclassical propagator has permitted the first semiclassical vibrational spectroscopy calculations on real molecular systems for which high-level potential energy surfaces (PES) are available. However, the approach is strongly limited by the necessity to rely on precise, but fast-to-compute, versions of the potential. For large molecular and supra-molecular systems analytical potentials are usually not available because of the difficulty to fit an analytical expression to such large systems, or, in case a force field exists, it may be not accurate enough for reliable spectroscopic calculations. A way to overcome this PES issue is to employ ab initio on-the-fly (i.e., direct) molecular dynamics simulations, in which the potential energy and gradient calls are performed at each time step while the dynamics is in progress [38]. These simulations are very computationally demanding though, and the number of trajectories required, even by a time averaged semiclassical simulation, is too large and not affordable.

Nevertheless, ab initio on-the-fly semiclassical spectroscopy has become possible by the recent introduction of the multiple coherent (MC) technique [39]. The foundation of MC SCIVR lies in pioneering work by De Leon and Heller [40] who demonstrated on low-dimensional model systems that accurate semiclassical eigenenergies and eigenfunctions can be obtained even by means of a single trajectory if it is run at the correct (quantum) energy. MC SCIVR assumes that reliable frequency estimates can be secured also for real molecular systems if a single trajectory is employed with an energy in the neighborhood of the true (but unknown) quantum one. A straightforward way to accomplish this goal is to work in normal mode coordinates by starting the trajectory from the equilibrium molecular geometry ($\vec{q}_{eq}$) with momenta selected according to a harmonic approximation, i.e., $\vec{p}_{eq} = \sqrt{(2\vec{n}+1)\hbar\vec{\omega}}$. Another pillar on which MC SCIVR is based concerns the choice of the reference state $|\Xi\rangle$. In order to enhance the Fourier transform signal of the vibrational mode of

interest, the reference state is chosen as a coherent state centered at the phase-space point $(\vec{p}_{\rm eq}, \vec{q}_{\rm eq})$ from which the classical trajectory has originated. In general, one different trajectory per each of the N_{st} states to be investigated is run. For each state (K), the corresponding reference state $|\Xi^{(K)}\rangle$ is appropriately chosen as

$$|\Xi^{(K)}\rangle = \prod_{J=1}^{N_{vib}} \sum_{\alpha=1}^{N_\alpha} \varepsilon_{\alpha,J}^{(K)} |\, p_{\mathrm{eq},\alpha,J}^{(K)},\ q_{\mathrm{eq},\alpha,J}^{(K)}\rangle. \quad K = 1, \ldots, N_{st}. \tag{19.44}$$

N_α is the number of coherent states and associated coefficients $\varepsilon_{\alpha,J}^{(K)}$ that are employed to enforce parity or molecular symmetry to favor detection of signals corresponding to specific mode excitations or symmetry species. The examples reported below will clarify this aspect. Equation (19.43) still serves as the working formula and, for the exact case of the harmonic oscillator, the phase-space integration is substituted by a sum over the tailored trajectories $(N_{tr} = N_{st})$ [41–43]

$$I(E) = \left(\frac{1}{2\pi\hbar}\right)^{N_{vib}} \sum_{K=1}^{N_{tr}} \frac{1}{2\pi\hbar T} \left| \int_0^T dt e^{\frac{i}{\hbar}[S_t^{(K)}(\vec{p}_0^{\,(K)}, \vec{q}_0^{\,(K)}) + Et + \phi_t^{(K)}]} \langle \Xi^{(K)} | \vec{p}_t^{\,(K)} \vec{q}_t^{\,(K)} \rangle \right|^2. \tag{19.45}$$

In the general case of an anharmonic and coupled potential, the total spectrum is obtained as the convolution of the spectral features provided by each trajectory in an energy range close to the energy of the trajectory.

The water molecule serves as a representative example of the technique. There are $N_{vib} = 3$ vibrational degrees of freedom, and initially we consider simulations based on just one classical trajectory $(N_{tr} = 1)$ to point out the features of a specific choice of the reference state $|\Xi\rangle$ [44]. By choosing $N_\alpha = 1$ and $|\Xi\rangle = \prod_{J=1}^{N_{vib}} \varepsilon_J |p_{\mathrm{eq},J}, q_{\mathrm{eq},J}\rangle$, the total power spectrum can be easily simulated. We introduce for shorthand notation the total vector $(\vec{\varepsilon})$ collecting the coefficients ε_J, which, in this case, is made of three elements (one per degree of freedom) all equal to 1. Figure 19.2 shows indeed the time averaged power spectrum obtained by running a single classical trajectory with harmonic zero point energy (indicated in Table 19.1 as "TA SCIVR 1traj (zpe)") and $\vec{\varepsilon} = (1;1;1)$ – semicolons separate the different degrees of freedom – on an analytical potential energy surface constructed

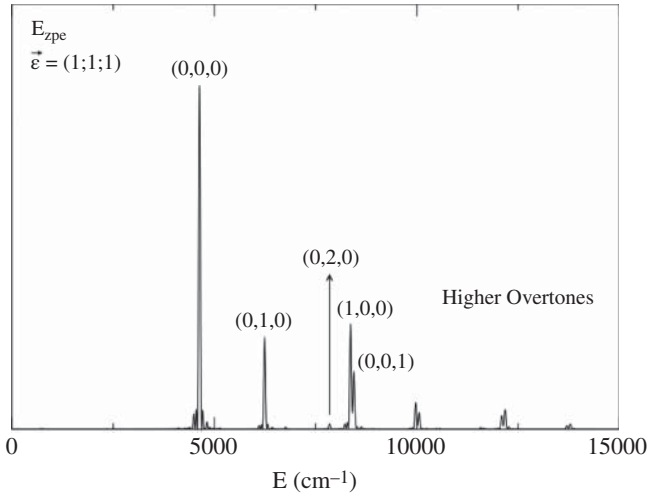


Figure 19.2 Total power spectrum of water using a single trajectory with harmonic zero point energy. Peaks are assigned and labeled by means of the usual quantum harmonic notation (symmetric stretch, bending, asymmetric stretch).

Table 19.1 Values (cm^{-1}) of zero point energy and some vibrational transitions of water. Frequency estimates are given in harmonic approximation (second column), with time averaged semiclassical dynamics based on a single trajectory with harmonic zero-point energy (third column), by means of MC SCIVR (fourth column), and quantum mechanically with a DVR technique (last column). In the last row the mean absolute error with respect to the QM values is reported. The usual quantum harmonic state notation has been employed with the first index corresponding to the symmetric stretch, the second index to the bending mode, and the third index to the asymmetric stretch.

Transition	Harm	TA SCIVR 1traj (zpe)	MC SCIVR	QM (DVR)
(0,0,0)	4711	4632	4632	4660
(0,0,0) → (0,1,0)	1650	1608	1584	1587
(0,0,0) → (0,2,0)	3300	3209	3171	3139
(0,0,0) → (1,0,0)	3830	3732	3706	3716
(0,0,0) → (0,0,1)	3941	3813	3813	3803
(0,0,0) → (1,1,0)	5480	5340	5231	5292
(0,0,0) → (0,1,1)	5591	5423	5307	5350
(0,0,0) → (2,0,0)	7660	7461	7410	7417
(0,0,0) → (0,0,2)	7882	7545	7500	7499
MAE	176	40	19	–

by Thiel et al. [45]. Due to the energetics of the trajectory, the most accurate peak is expected to be the zero-point energy (ZPE) one.

To assign peaks in the power spectrum with more confidence it is possible to insert symmetry into the calculations. This is obtained by duplicating the number of coherent states ($N_\alpha = 2$) that make up the reference state in Eq. (19.44). The dimensionality of $\vec{\varepsilon}$ is also doubled. Specifically, parity symmetry can be enforced by choosing the reference states as $|\Xi\rangle = \prod_{J=1}^{N_{vib}}(\varepsilon_{1,J}|p_{\mathrm{eq},J}, q_{\mathrm{eq},J}\rangle + \varepsilon_{2,J}|-p_{\mathrm{eq},J}, q_{\mathrm{eq},J}\rangle)$. In fact, the sign of $\varepsilon_{2,J}$ determines if an odd ($\varepsilon_{2,J} = -1$) or even ($\varepsilon_{2,J} = 1$) number of quanta in the mode J is associated to the peak to be assigned. For instance, for the ground state $\vec{\varepsilon} = (1, 1; 1, 1; 1, 1)$, while for the first excited state $\vec{\varepsilon} = (1, -1; 1, 1; 1, 1)$. This is also evident from the upper plot in Figure 19.3 where in the first of the four panels only the ZPE peak is enhanced. There are also minor peaks which can be barely seen at high energy. Some of them correspond to double excitations of the modes. In the other three panels of the same plot the first excited states of each mode have been enhanced with an appropriate choice of the ε components. All spectra were obtained from the very same single trajectory with harmonic zero point energy.

As for the molecular symmetry we note that symmetric stretch and bending of water are of a_1 symmetry, while the asymmetric stretch is of b_2 symmetry. This means that we can enforce the desired symmetry and get the corresponding peaks in the power spectrum by employing the reference state $|\Xi\rangle = \prod_{J=1}^{N_{vib}}(\lambda_{1,J}|p_{\mathrm{eq},J}, q_{\mathrm{eq},J}\rangle + \lambda_{2,J}|-p_{\mathrm{eq},J}, -q_{\mathrm{eq},J}\rangle)$. Note that the second coherent state in the linear combination is now centered at $(-p_{\mathrm{eq},J}, -q_{\mathrm{eq},J})$, and the symmetry parameters have been renamed as λ to help the reader. To discriminate between peaks, we set the value of the λ_2 component relative to the asymmetric stretch equal to 1 (a_1 symmetry) or -1 (b_2 symmetry). Figure 19.3 reports on the bottom plot the two simulations for water in which peaks are distinguished according to their molecular symmetry. The proposed methodology to enforce parity or molecular symmetry is rigorous for the harmonic oscillator but, as demonstrated by the simulations here presented, it is effective also for realistic molecular systems.

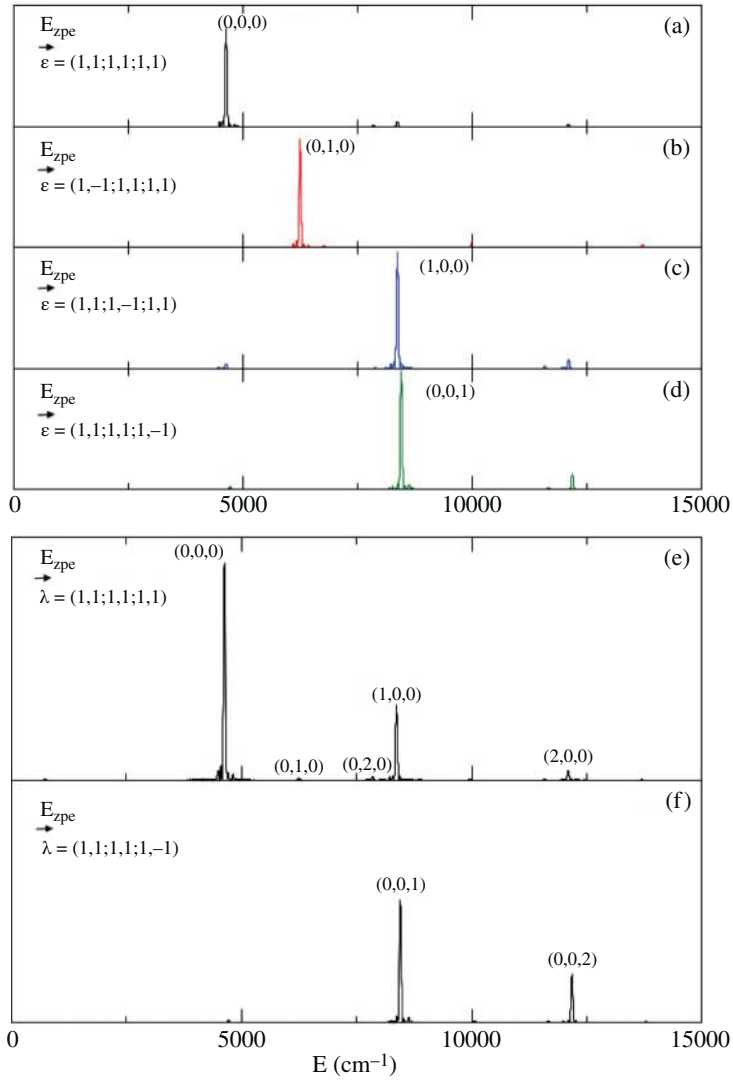


Figure 19.3 Top: Selection of water semiclassical eigenvalues from a single trajectory with harmonic zero point energy. Panel (a): zero point energy (4632 cm^{-1}, black line); panel (b): bending mode (red); panel (c): symmetric stretch (blue); panel (d): asymmetric stretch (green). Frequencies of the fundamental transitions can be obtained as differences with respect to the zero point energy. They are: $\omega_b = 1608$ cm^{-1}; $\omega_s = 3732$ cm^{-1}; $\omega_a = 3813$ cm^{-1}. Bottom: peaks are selected on the basis of molecular symmetry. On the top panel (e) modes with a_1 symmetry are presented, while the bottom panel (f) illustrates peaks with b_2 symmetry. Note that in the simulations the elements of $\vec{\varepsilon}$ and $\vec{\lambda}$ are ordered according to the ascending harmonic frequency of the modes.

So far a single harmonic zpe trajectory has been employed but, in the true spirit of the multiple coherent states technique ($N_{tr} = N_{st}$), results can be refined in a peak-by-peak fashion by running for each mode one trajectory with one quantum of excitation in that specific mode and employing the corresponding reference state. This leads to a better estimate of fundamental transition frequencies and especially overtones. Table 19.1 reports a comparison of the frequencies provided by different approaches. As expected, the enhanced accuracy of a multiple coherent technique is

evident when comparing semiclassical estimates to the quantum benchmark obtained by means of a discrete variable representation (DVR) calculation in normal mode coordinates. We stress that the simulations yield the spectrum (i.e., the eigenvalues) of the vibrational Hamiltonian, so we are able to assign a frequency even to those vibrational transitions which are not IR or Raman active.

Among remarkable applications of the MC-SCIVR method we recall a study of ammonia, which has permitted detection of the spectrum of this molecule and to mimic its peculiar tunneling splitting feature with just eight trajectories [46], and an investigation of the simplest aminoacid: glycine [47]. Glycine is characterized by several conformers as demonstrated by the multiple minimum structure of its potential energy surface. An ab initio on-the-fly MC-SCIVR approach has allowed investigation of all these conformers in full dimensionality and to estimate the potential effect of conformer interconversion on vibrational frequencies. More details on this system will be presented in Section 19.10. The advantage with respect to other methods based on calculations confined to a single well is evident. Furthermore, electronic theory calculations are needed only for the geometries experienced along the dynamics so a global full-dimensional surface (which can be very difficult to construct) is not required.

19.7 The "Divide-and-Conquer" SCIVR

Perhaps the most relevant issue that semiclassical dynamics and other quantum approaches must face is the scaling of computational overheads with the dimensionality of the system under investigation. In the case of SC approaches the challenge is to get a well resolved spectroscopic signal when the number of degrees of freedom exceeds 25–30. The principal reason for this issue is the multidimensional coherent state overlap which characterizes the SC formulation as in the case, for instance, of Eq. (19.43). In fact, a sensible signal requires all mono-dimensional coherent state overlaps to be simultaneously not negligible at each step of the dynamics, a request which is harder and harder to satisfy as the dimensionality of the problem increases. There is an evident difference with respect to classical simulations based on the Fourier transform of the dipole–dipole autocorrelation function in which a scalar product of 3-dimensional vectors is involved whatever the dimensionality of the system is.

Approaches have been developed to identify the effective vibrational modes for the calculations [48], while here we focus on the possibility to work in reduced dimensionality by projecting the full-dimensional problem onto a set of lower dimensional ones, where it is easier to have a recurring coherent state overlap. The total spectrum is eventually recollected as a convolution of the lower dimensional spectra [49]. Figure 19.4 depicts this "divide-and-conquer" (DC-SCIVR) idea. We note that the trajectory, i.e., the classical dynamics, is still performed in full dimensionality,

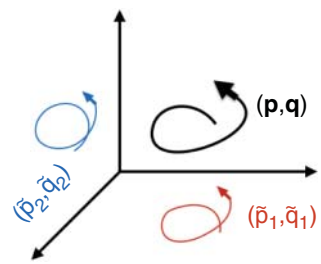


Figure 19.4 Pictorial representation of the DC-SCIVR idea. The black line is the full-dimensional trajectory while the red and blue lines represent its projection onto two different subspaces.

while associated quantities are projected onto different subspaces in which the semiclassical simulations are undertaken. The multiple coherent states reduced-dimensionality (MC-DC SCIVR) working formula becomes

$$\tilde{I}(E) = \left(\frac{1}{2\pi\hbar}\right)^{M} \sum_{K=1}^{N_{tr}} \frac{1}{2\pi\hbar T} \left| \int_0^T dt e^{\frac{i}{\hbar}[\tilde{S}_t^{(K)}(\tilde{\vec{p}}_0^{(K)}, \tilde{\vec{q}}_0^{(K)}) + Et + \tilde{\phi}_t^{(K)}]} \langle \tilde{\Xi}^{(K)} | \tilde{\vec{p}}_t^{(K)} \tilde{\vec{q}}_t^{(K)} \rangle \right|^2, \tag{19.46}$$

where M is the dimensionality of the subspace. Equation (19.46) is at the heart of the MC-DC-SCIVR approach. Mathematically, the projection process is equivalent to a singular value decomposition which allows to reduce the dimensionality of the system.

In practice, after the full-dimensional trajectory has been run, the reduced quantities appearing in Eq. (19.46) are calculated. Specifically, the prefactor phase is estimated by employing the appropriate sub-block of monodromy matrix elements, and evaluation of the coherent state overlap is straightforward. Calculation of the projected action is more complicated instead, due to the non-separability of the potential energy. In fact, if the degrees of freedom not belonging to the M-dimensional subspace are simply downgraded to parameters, then the resulting potential $V(\vec{q}(t)) \equiv V(\vec{q}_M(t); \vec{q}^{\,\mathrm{eq}}_{N_{vib}-M})$ is in general not suitable to reproduce the "correct" low-dimensional dynamics. An ideal reduced-dimensional potential should permit generation of a reduced-dimensional dynamics which, starting from the initial point in phase space $(\vec{p}_0, \vec{q}_0)$, visits all points $(\vec{p}_t, \vec{q}_t)$ obtained projecting the full-dimensional trajectory onto the subspace. This is not the case for the trivial definition proposed above. Instead, it is possible to introduce a time dependent external field

$$\begin{aligned} \tilde{V}(\vec{q}_M(t)) &= V(\vec{q}_M(t); \vec{q}^{\,\mathrm{eq}}_{N_{vib}-M}) + \lambda(t); \\ \lambda(t) &= V(\vec{q}_M(t); \vec{q}_{N_{vib}-M}(t)) - V(\vec{q}_M(t); \vec{q}^{\,\mathrm{eq}}_{N_{vib}-M}) - V(q_M^{\mathrm{eq}}; \vec{q}_{N_{vib}-M}(t)). \end{aligned} \tag{19.47}$$

This choice of the field returns the exact dynamics in the case of a separable potential, while it still provides a reliable approximation in the general instance of a non-separable potential. This means that in the DC-SCIVR simulations the projected action must be calculated by means of the potential defined in Eq. (19.47). This is achieved by performing, after the trajectory is complete, a single-energy calculation for each configuration of the dynamics upon substitution of the coordinates of the modes belonging to the subspace of interest with their equilibrium values. As usual, the transformation from normal mode coordinates to Cartesian ones is done by means of the matrix of eigenvectors of the equilibrium Hessian. The interested reader can find further details in Refs. [49–51].

Once the mathematical formalism of DC-SCIVR has been introduced, an efficient strategy to partition the full-dimensional problem into lower-dimensional ones must be devised. An educated choice of subspaces has to deal with the trade-off between simulation accuracy, which is maximum (at the net of compensation of errors) for the full-dimensional system, and the necessity to work in reduced dimensionality to get a sensible spectroscopic signal from the Fourier transform. The intuitive way to proceed consists in collecting the normal modes that are more strongly interacting into the same subspace. For this purpose a few strategies have been developed. The first one (known as the Hessian decomposition method) is based on averaging the off-diagonal elements of the normal mode Hessian matrix (which are initially 0 since the normal-mode Hessian is diagonal at the starting equilibrium geometry) dynamically along the harmonic zero-point-energy trajectory and then comparing them to an arbitrary threshold. The average Hessian elements are taken as an estimate of the coupling between the modes, so, if they are above the threshold, the two involved modes are set into the same subspace. The method has been proved effective in a number of applications [49, 50] even if it suffers from the arbitrariness of the threshold choice. In fact, if the threshold is

Threshold | **Averaged Hessian** | **Subspaces**

$\varepsilon_{thr} = 1.5*10^{-6}$

Hess	V_1	V_2	V_3	V_4	V_5	V_6
V_1		$1.8*10^{-6}$	$2.5*10^{-5}$	$8.2*10^{-7}$	$3.3*10^{-6}$	$7.1*10^{-8}$
V_2	$1.8*10^{-6}$		$1.3*10^{-5}$	$6.6*10^{-6}$	$2.0*10^{-6}$	$7.5*10^{-8}$
V_3	$2.5*10^{-5}$	$1.3*10^{-5}$		$4.8*10^{-7}$	$9.1*10^{-7}$	$3.6*10^{-9}$
V_4	$8.2*10^{-7}$	$6.6*10^{-6}$	$4.8*10^{-7}$		$8.8*10^{-8}$	$1.7*10^{-4}$
V_5	$3.3*10^{-6}$	$2.0*10^{-6}$	$9.1*10^{-7}$	$8.8*10^{-8}$		$5.2*10^{-7}$
V_6	$7.1*10^{-8}$	$7.5*10^{-8}$	$3.6*10^{-9}$	$1.7*10^{-4}$	$5.2*10^{-7}$	

Sub I: V_1, V_2, V_3, V_5

Sub II: V_4, V_6

$\varepsilon_{thr} = 1.5*10^{-5}$

Hess	V_1	V_2	V_3	V_4	V_5	V_6
V_1		$1.8*10^{-6}$	$2.5*10^{-5}$	$8.2*10^{-7}$	$3.3*10^{-6}$	$7.1*10^{-8}$
V_2	$1.8*10^{-6}$		$1.3*10^{-5}$	$6.6*10^{-6}$	$2.0*10^{-6}$	$7.5*10^{-8}$
V_3	$2.5*10^{-5}$	$1.3*10^{-5}$		$4.8*10^{-7}$	$9.1*10^{-7}$	$3.6*10^{-9}$
V_4	$8.2*10^{-7}$	$6.6*10^{-6}$	$4.8*10^{-7}$		$8.8*10^{-8}$	$1.7*10^{-4}$
V_5	$3.3*10^{-6}$	$2.0*10^{-6}$	$9.1*10^{-7}$	$8.8*10^{-8}$		$5.2*10^{-7}$
V_6	$7.1*10^{-8}$	$7.5*10^{-8}$	$3.6*10^{-9}$	$1.7*10^{-4}$	$5.2*10^{-7}$	

Sub I: V_1, V_3

Sub II: V_4, V_6

Sub III: V_2

Sub IV: V_5

Figure 19.5 Influence of the choice of the threshold parameter on the partition of a 6-dimensional space into subspaces. For $\varepsilon = 1.5 \cdot 10^{-6}$ the six vibrational modes are collected into a 4-dimensional space and a 2-dimensional one. For $\varepsilon = 1.5 \cdot 10^{-5}$ there are two bi-dimensional subspaces and two mono-dimensional ones identified by a red entry in the diagonal. Diagonal Hessian elements are not of interest and they have not been reported.

too small, then the full-dimensional work space is left unchanged, whereas, if the threshold is too big, then all subspaces are inaccurately chosen to be mono-dimensional ones. Figure 19.5 exemplifies the different partition into subspaces obtained for a 6-mode model system when modifying the threshold parameter. Note that in the top panel modes ν_3 and ν_5 are in the same subspace even if they are not sufficiently coupled. This is because they are both strongly coupled to other modes (ν_1 and ν_2 in this specific case). A similar methodology, but still dependent on an arbitrary parameter, is based on an average along the trajectory of the monodromy matrix elements instead of the Hessian [48].

To make the subspace partition less arbitrary, a procedure based on the unitarity of the determinant of the phase-space Jacobian in the subspaces has been developed [50]. This approach is known as the Jacobi decomposition method. The starting point is the definition of the full dimensional Jacobian (equivalent to the monodromy matrix) as

$$\mathbf{J}(t) = \begin{pmatrix} \partial\vec{q}_t/\partial\vec{q}_0 & \partial\vec{q}_t/\partial\vec{p}_0 \\ \partial\vec{p}_t/\partial\vec{q}_0 & \partial\vec{p}_t/\partial\vec{p}_0 \end{pmatrix}, \tag{19.48}$$

whose determinant is equal to 1 at all times. This property is easy to demonstrate because at time 0 the determinant equals 1, and its time derivative is identically zero independently of time. A direct and important consequence is that $d\vec{p}_t d\vec{q}_t = d\vec{p}_0 d\vec{q}_0$, one of the ways to express Liouville's theorem of Hamiltonian mechanics. If we define a set of N_{sub} subspaces as requested by DC SCIVR, the determinant of the full-dimensional Jacobian can be calculated as the product of the partial determinants defined by the projected positions and momenta, i.e., $det(\mathbf{J}) = \prod_{i=1}^{N_{sub}} det(\tilde{\mathbf{J}}_i)$. If the

system is separable, the unitarity of the determinant is valid also in the reduced-dimensionality subspaces (i.e., $d\vec{p}_t d\vec{q}_t = d\vec{p}_0 d\vec{q}_0$) but this is no longer true for the generic case of a non-separable system. In general, the determinant of the Jacobian is not conserved in the subspaces, so a good criterion to select the best way to collect vibrational modes is one that preserves as much as possible the Jacobian determinant. For this purpose it is possible to employ an iterative procedure consisting in evaluating along the dynamics the Jacobian determinants of all possible reduced-dimensional subspaces. At each time step and for each dimensionality, the best subspace is the one with the closest Jacobian determinant to unity. Within a given dimensionality the overall best subspace is then chosen to be the one that most frequently had $det(\tilde{\mathbf{J}})$ closest to unity. Finally, the representative candidates of all subspaces of different dimensionality are compared among them and the one with the Jacobian determinant closest to unity is selected. The whole procedure is then repeated for the remaining degrees of freedom until all of them have been assigned. Figure 19.6 sketches the methodology for a 6-dimensional model system. Note that in the example a 4-dimensional subspace is selected, which means the procedure will be repeated involving only modes four and six that end

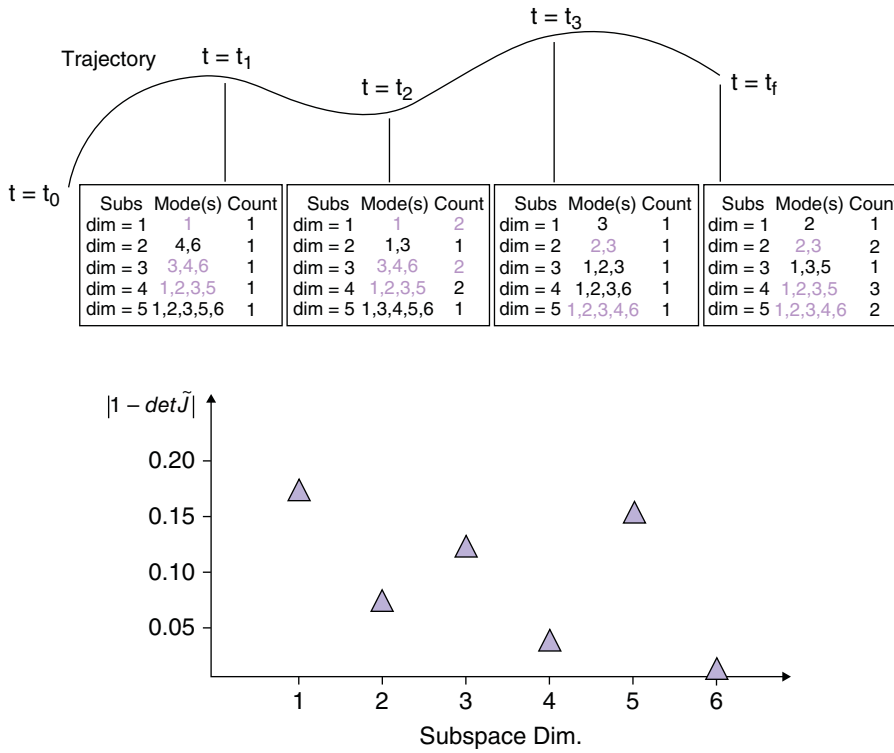


Figure 19.6 Schematic representation of the Jacobi decomposition method. First the best set of modes for each possible subspace dimensionality is calculated at every step along the trajectory. This is shown in the four boxes, where the "Subs" column indicates the subspace dimensionality and the "Count" one the incremental number of times that particular set of modes has been selected along the trajectory. For each subspace dimensionality the most frequent mode combination acts as the representative of that subspace dimensionality and it is highlighted (violet). Then, $|1 - \det\tilde{J}|$ values of all the representatives are averaged over the trajectory and compared (bottom, triangles). The 4-dimensional subspace made of modes 1,2,3,5 has the lowest $|1 - \det\tilde{J}|$ average value and it is therefore selected. Modes 4 and 6 are left and the procedure is reiterated for these two remaining degrees of freedom to check out if they belong to the same bi-dimensional subspace or if they split into a pair of mono-dimensional subspaces.

up either in a single 2-dimensional subspace or in two separate mono-dimensional subspaces. The latter instance is, in principle, not desirable but part of the interactions are anyway preserved due to the full-dimensional trajectory on which projected quantities are based. The Hessian and Jacobi decomposition methods provide generally different types of subspace partitions. On the one hand, the Jacobi approach has been shown to lead to more accurate results [50], but it is harder and harder to apply it as the total number of degrees of freedom increases due to the necessity to check all possible vibrational mode subsets. On the other hand, the Hessian decomposition approach is less accurate and depends on an arbitrary parameter, but its application is instantaneous once the averaged Hessian has been computed.

We conclude this section by reporting on an application of DC SCIVR to C_{60} fullerene, a system made of 174 vibrational degrees of freedom. A force field originally created to study graphene layers is adapted to permit analytical potential calls but, nevertheless, these are very time consuming, so a MC-DC-SCIVR approach is employed by running 175 classical trajectories about 1.2 ps long (1 at harmonic zero point energy plus other 174, each one with a harmonic quantum of excitation in a different mode). The subspaces are determined by means of the Hessian decomposition criterion based on the harmonic zero-point energy. We first scan the maximum subspace dimensionality for different values of the threshold parameter as shown in Figure 19.7, and then choose for the threshold the value $\varepsilon = 10^{-6}$ corresponding to a maximum subspace dimensionality of 25. This choice is driven by the trade-off between expected accuracy and the possibility to collect a sensible spectral signal. The whole degrees of freedom are consequently collected into one

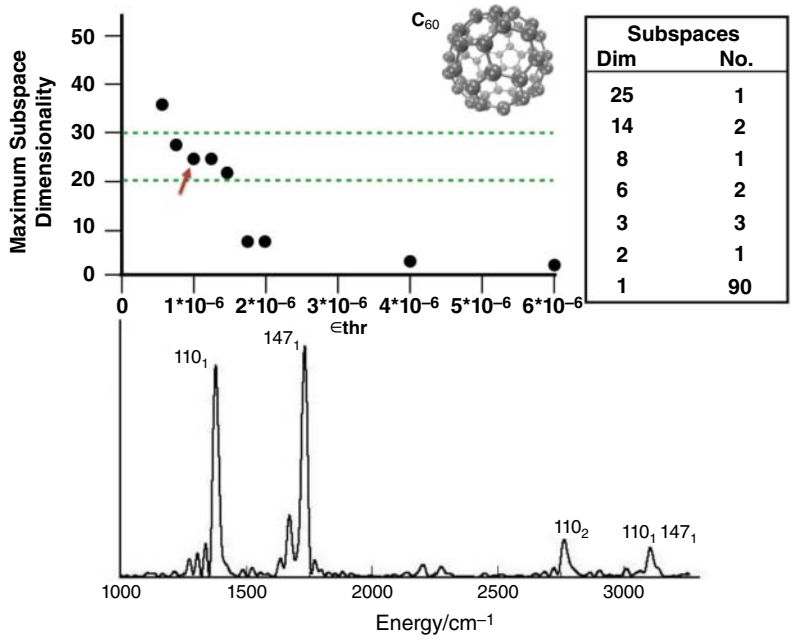


Figure 19.7 Maximum subspace dimensionality versus threshold when applying the Hessian decomposition method to the C_{60} fullerene molecule (top). The red arrow indicates the chosen threshold and the corresponding maximum subspace dimensionality. The green dashed lines determine the desired range for the maximum subspace dimensionality. The table reports the detailed partition into subspaces. On the bottom, a bi-dimensional DC-SCIVR spectrum (involving modes 110 and 147) is shown (the zpe of the subspace has been shifted to zero).

25-dimensional, two 14-dimensional, one 8-dimensional, two 6-dimensional, three 3-dimensional, one bi-dimensional, and 90 mono-dimensional subspaces.

Once the subspaces have been determined, MC-DC-SCIVR calculations can be performed. On the bottom of Figure 19.7 an example concerning a bi-dimensional subspace is reported where the two fundamentals and some overtones are clearly detected. For detailed results we refer to a previous publication (see Ref. [49]). Here we conclude this section by pointing out that even if, as already demonstrated, molecular symmetry can be incorporated into semiclassical dynamics to facilitate peak assignment, it does not help speed up calculations and it is not necessary for a good outcome of DC-SCIVR simulations, which are successful independently of symmetry.

19.8 Mixed SCIVR Dynamics: Towards Semiclassical Spectroscopy in Condensed Phase

The various semiclassical methodologies so far presented have in common the characteristic to be based on the evolution of Gaussian wave packets of fixed width. Such propagators go under the collective name of "Frozen Gaussian propagators". Another family of semiclassical Gaussian propagators has been introduced [52] in which the Gaussian width is permitted to change in time. They are called "Thawed Gaussian propagators". In practice (for simplicity we illustrate equations in one dimension, but generalization to multiple dimensions is straightforward) the wave packet is chosen to be of the coherent state form

$$\Xi(q,t) = \left(\frac{\Gamma_0}{\pi}\right)^{1/4} \exp\left[-\frac{\Gamma_t}{2}(q-q_t)^2 + \frac{\mathrm{i}}{\hbar}p_t(q-q_t) + \frac{\mathrm{i}}{\hbar}\delta_t\right], \tag{19.49}$$

with the usual Hamiltonian evolution for q_t and p_t, and the appropriate equations of motions for Γ_t and δ_t

$$\begin{aligned} -\mathrm{i}\hbar\frac{d\Gamma_t}{dt} &= -\hbar^2\Gamma_t^2 + \frac{d^2V(q_t,t)}{dq_t^2} \\ \frac{d\delta_t}{dt} &= \frac{p_t^2}{2} - V(q_t,t) - \frac{\hbar^2}{2}\Gamma_t. \end{aligned} \tag{19.50}$$

Thawed Gaussian propagation is known to be less accurate than the Heller–Herman–Kluk–Kay one, but its major flexibility allows results of good accuracy to be gained with just a single Gaussian propagation.

An interesting development of semiclassical dynamics involving both Frozen and Thawed Gaussian propagators is represented by Grossmann's hybrid (or mixed) SCIVR and its recent simplified version [53]. The method is based on the observation that the Thawed Gaussian propagator can be obtained as an approximation to the HHKK one [54]. In mixed SCIVR, the higher HHKK level of theory is reserved for a few degrees of freedom, while all the others are treated by means of the computationally cheaper Thawed Gaussian propagator. A promising application of the technique is represented by the spectroscopic investigation of a molecular system embedded in an environment. In spectroscopic applications of the mixed SCIVR the starting point is the time-averaged version of the HHKK propagator in Eq. (19.43). On this propagator the mixed approximation is inserted. The total $2F$ phase-space variables are divided into $2F_{\mathrm{HHKK}}$ for the system and $2F_{\mathrm{TG}}$ for the bath. The system, on which more accurate information is sought, is represented with the HHKK label, meaning it is treated at the HHKK level of theory, whereas the environmental degrees of freedom are indicated with the subscript TG and treated by means of the Thawed Gaussian propagator. The

method, similarly to DC SCIVR, is characterized by classical dynamics runs in full dimensionality. The partition of the degrees of freedom is adopted only in the semiclassical formalism. In fact, if the Gaussian reference state $|\Xi\rangle = |\vec{p}_{\rm eq}, \vec{q}_{\rm eq}\rangle$ is chosen, where $\vec{q}_{\rm eq}$ is the equilibrium position and $\vec{p}_{\rm eq}$ is the momentum corresponding to an approximate eigenenergy, then the initial phase space coordinates $(\vec{p}_0, \vec{q}_0)$ are defined as

$$\vec{p}_0 = \begin{pmatrix} \vec{p}_{\rm HHKK,0} \\ \vec{p}_{\rm eq,TG} \end{pmatrix}, \qquad \vec{q}_0 = \begin{pmatrix} \vec{q}_{\rm HHKK,0} \\ \vec{q}_{\rm eq,TG} \end{pmatrix}. \tag{19.51}$$

The HHKK initial conditions $(\vec{p}_{\rm HHKK,0}, \vec{q}_{\rm HHKK,0})$ are selected by means of the usual Monte Carlo sampling around $(\vec{p}_{\rm eq,HHKK}, \vec{q}_{\rm eq,HHKK})$, while the bath starting coordinates are always at the equilibrium positions, $(\vec{p}_{\rm TG,0}, \vec{q}_{\rm TG,0}) = (\vec{p}_{\rm eq,TG}, \vec{q}_{\rm eq,TG})$. Then, expansion of classical trajectories and action to first and second order, respectively, in the displacement coordinates of the bath subspace allows the phase-space integration over the original initial bath conditions $(\vec{p}_{\rm TG,0}, \vec{q}_{\rm TG,0})$ to be performed analytically. In this way the dimensionality of the phase space integration is reduced. Finally, by expanding all quantities to second order for the bath variables we arrive at the separable mixed TA-SCIVR (M-TA SCIVR)

$$\begin{aligned} I(E) = \frac{1}{(2\hbar)^F} \frac{1}{\pi^{F_{\rm HHKK}}} \frac{1}{2\pi\hbar T} \int {\rm d}\vec{p}_{\rm HHKK,0} \int {\rm d}\vec{q}_{\rm HHKK,0} \left| \int_0^T {\rm d}t\, {\rm e}^{{\rm i}[Et + \phi_t(\vec{p}_0,\vec{q}_0) + S_t(\vec{p}_0,\vec{q}_0)]/\hbar} \right. \\ \times \langle \vec{p}_{\rm eq,HHKK}, \vec{q}_{\rm eq,HHKK} | \vec{p}_{{\rm HHKK},t}, \vec{q}_{{\rm HHKK},t} \rangle \langle \vec{p}_{\rm eq,TG}, \vec{q}_{\rm eq,TG} | \vec{p}_{{\rm TG},t}, \vec{q}_{{\rm TG},t} \rangle \\ \left. \times \frac{1}{[\det(\mathbf{A}(t) + \mathbf{A}^*(t))]^{1/4}} \exp\left\{ \frac{1}{4} \vec{b}_t^{\,\rm T} (\mathbf{A}(t) + \mathbf{A}^*(t))^{-1} \vec{b}_t \right\} \right|^2 . \end{aligned} \tag{19.52}$$

The elements of the matrix $\mathbf{A}(t)$ are defined as

$$\begin{aligned} A_{11}(t) &= \frac{1}{4} M_{21}^{\rm T}(t)\, \mathbf{\Gamma} M_{21}(t) + \frac{1}{4\hbar^2} M_{11}^{\rm T}(t)\, \mathbf{\Gamma}^{-1} M_{11}(t) \\ A_{12}(t) &= \frac{1}{4} M_{21}^{\rm T}(t)\, \mathbf{\Gamma} M_{22}(t) + \frac{1}{4\hbar^2} M_{11}^{\rm T}(t)\, \mathbf{\Gamma}^{-1} M_{12}(t) \\ A_{21}(t) &= \frac{1}{4} M_{22}^{\rm T}(t)\, \mathbf{\Gamma} M_{21}(t) + \frac{1}{4\hbar^2} M_{12}^{\rm T}(t)\, \mathbf{\Gamma}^{-1} M_{11}(t) \\ A_{22}(t) &= \frac{1}{4} M_{22}^{\rm T}(t)\, \mathbf{\Gamma} M_{22}(t) + \frac{1}{4\hbar^2} M_{12}^{\rm T}(t)\, \mathbf{\Gamma}^{-1} M_{12}(t), \end{aligned} \tag{19.53}$$

while the vector $\vec{b}(t) \equiv (\vec{b}_{1,t}^{\,\rm T}, \vec{b}_{2,t}^{\,\rm T})^{\rm T}$ is made of the sub-vectors

$$\begin{aligned} \vec{b}_{1,t}^{\,\rm T} &= -\frac{1}{2} (\vec{q}(t) - \vec{q}(0))^{\rm T} \left[\mathbf{\Gamma} M_{21}(t) + \frac{\rm i}{\hbar} M_{11}(t) \right] \\ &\quad - \frac{1}{2\hbar^2} (\vec{p}(t) - \vec{p}(0))^{\rm T} \left[\mathbf{\Gamma}^{-1} M_{11}(t) - {\rm i}\hbar M_{21}(t) \right] \\ \vec{b}_{2,t}^{\,\rm T} &= -\frac{1}{2} (\vec{q}(t) - \vec{q}(0))^{\rm T} \left[\mathbf{\Gamma} M_{22}(t) + \frac{\rm i}{\hbar} M_{12}(t) \right] \\ &\quad - \frac{1}{2\hbar^2} (\vec{p}(t) - \vec{p}(0))^{\rm T} [\mathbf{\Gamma}^{-1} M_{12}(t) - {\rm i}\hbar M_{22}(t)], \end{aligned} \tag{19.54}$$

where $(\vec{p}(t), \vec{q}(t))$ is the trajectory starting from the initial conditions defined in Eq. (19.51). The M_{ij} in the above equations are the $F \times F_{\mathrm{TG}}$ sub-matrices of the stability matrix,

$$\begin{aligned} M_{11}(t) &= \frac{\partial \vec{p}_t}{\partial \vec{p}_{\mathrm{TG},0}}, \quad M_{12}(t) = \frac{\partial \vec{p}_t}{\partial \vec{q}_{\mathrm{TG},0}}, \\ M_{21}(t) &= \frac{\partial \vec{q}_t}{\partial \vec{p}_{\mathrm{TG},0}}, \quad M_{22}(t) = \frac{\partial \vec{q}_t}{\partial \vec{q}_{\mathrm{TG},0}}. \end{aligned} \tag{19.55}$$

In Eq. (19.52) the Monte Carlo integration involves only the system coordinates, while the bath degrees of freedom are included through the dynamics which starts from the conditions in Eq. (19.51). The interested reader will find more details in Refs. [53, 55, 56].

The mixed SC-IVR approach is effective, but, when studying systems characterized by a large number of degrees of freedom, spectra are noisy and difficult to interpret. To overcome this issue a simplified approach to M-TA SCIVR (SAM) has been introduced. It is based on two assumptions. First, the following bath components are replaced by their analytic harmonic oscillator results

$$\begin{aligned} &\frac{1}{[\det(\mathbf{A}_{\mathrm{HO}}(t) + \mathbf{A}^*_{\mathrm{HO}}(t))]^{1/4}} \approx (2\hbar)^{F_{\mathrm{TG}}/2} \\ &\vec{b}^{\,\mathrm{T}}_{t,\mathrm{HO}}(\mathbf{A}_{\mathrm{HO}}(t) + \mathbf{A}^*_{\mathrm{HO}}(t))^{-1}\vec{b}_{t,\mathrm{HO}} \approx 0, \end{aligned} \tag{19.56}$$

(where HO stands for Harmonic Oscillator) which permits Eq. (19.52) to be considerably simplified, transforming it into

$$\begin{aligned} I(E) = &\frac{1}{(2\pi\hbar)^{F_{\mathrm{HHKK}}}} \frac{1}{2\pi\hbar T} \int \mathrm{d}\vec{p}_{\mathrm{HHKK},0} \int \mathrm{d}\vec{q}_{\mathrm{HHKK},0} \\ &\times \left| \int_0^T \mathrm{d}t \, \mathrm{e}^{\mathrm{i}[Et + \phi_t(\vec{p}_0,\vec{q}_0) + S_t(\vec{p}_0,\vec{q}_0)]/\hbar} \langle \vec{p}_{\mathrm{eq}}, \vec{q}_{\mathrm{eq}} | \vec{p}_t, \vec{q}_t \rangle \right|^2 . \end{aligned} \tag{19.57}$$

Then, the multiple coherent technique is employed and the final SAM working formula is

$$\begin{aligned} I(E) = &\frac{1}{(2\pi\hbar)^{F_{\mathrm{HHKK}}}} \frac{1}{2\pi\hbar T} \int \mathrm{d}\vec{p}_{\mathrm{HHKK},0} \int \mathrm{d}\vec{q}_{\mathrm{HHKK},0} \\ &\left| \int \mathrm{d}t \, \mathrm{e}^{\mathrm{i}[Et + \phi_t(\vec{p}_0,\vec{q}_0) + S_t(\vec{p}_0,\vec{q}_0)]/\hbar} \langle \vec{p}_{\mathrm{eq,HHKK}}, \vec{q}_{\mathrm{eq,HHKK}} | \vec{p}_{\mathrm{HHKK},t}, \vec{q}_{\mathrm{HHKK},t} \rangle \right|^2 . \end{aligned} \tag{19.58}$$

It should be pointed out that, in spite of the different derivations of SAM and DC SCIVR, both methods are in fact related. Adoption of the few-trajectory approach from MC SCIVR to Eq. (19.58) by replacing the HK initial state $|\vec{p}_{\mathrm{eq,\,HHKK}}, \vec{q}_{\mathrm{eq,\,HHKK}}\rangle$ with $|\Xi\rangle$ from Eq. (19.44) turns the phase space integration into a sum over (few) trajectories. The resulting equation has the same structure as the MC-DC-SCIVR working formula of Eq. (19.46), except for the specific form of the action and the prefactor phase. In this sense, SAM can be seen as an intermediate step between HK SCIVR and DC SCIVR. A remarkable application of the method has permitted a full study of anharmonic vibrations of the iodine molecule to be performed (I_2) in a krypton matrix made of 218 atoms [56]. This shows that SAM is a promising tool for performing highly accurate condensed phase spectroscopy in the near future.

19.9 Semiclassical Spectroscopy Workflow

Before showing some applications of SC dynamics, we present a workflow chart (Figure 19.8) summarizing the steps on which the various semiclassical approaches previously described are based. This chart keeps evolving as new semiclassical methodologies are being developed.

Relevant input data for the simulations include the reference state and the number of classical trajectories to be run. A preliminary routine optimization of the equilibrium geometry and calculation of harmonic frequencies (as usual via Hessian diagonalization) are requested. Afterwards, the chart branches according to the number of trajectories chosen. On the one hand, if just a single or a few trajectories are employed, then either thawed Gaussian propagation or the family of multiple coherent approaches is enabled. In this case, initial conditions for the trajectories are tailored, and upon classical evolution and Hessian calculation the corresponding mathematical expressions are used to get the spectrum. On the other hand, when Monte Carlo phase-space integration is performed by means of thousands of trajectories, initial conditions are sampled from a distribution (generally a Husimi one). Again, after dynamics evolution and Hessian calculation, application of the appropriate equations allows us to get the spectrum. A third instance is represented by M-TA SCIVR, which employs a multiple trajectory approach for the system and a thawed Gaussian (single trajectory) approach for the bath. For the required electronic structure calculations the freely available NWChem suite of codes is suggested [57] due to its interface with VENUS. VENUS is another free software for ab initio dynamical calculations [58] which includes semiclassical codes [59].

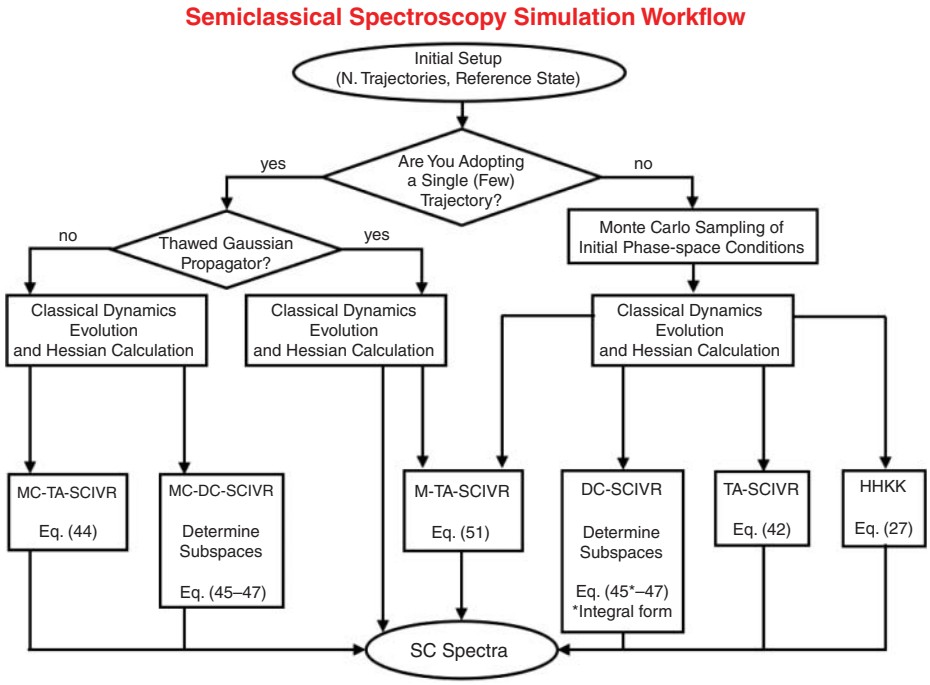


Figure 19.8 Workflow chart.

19.10 A Taste of Semiclassical Spectroscopy

A few relevant applications of semiclassical spectroscopy to molecular power, photoemission, and photoabsorption spectra are presented in this section. The semiclassical calculations of power and photoemission spectra involve electronic ground state dynamics, while photoabsorption spectra rely on electronic excited state dynamics in agreement with the formulae reported in the introductory section of the chapter.

The first application regards gas-phase glycine [47]. An analysis of the potential energy surface of this simple amino acid reveals quite a complex landscape with the presence of four conformers. Three of them are separated by small barriers which may favor conformer interconversion. The dynamics employed in the SC simulations shows that conformer interconversion takes place when exciting selectively some of the normal modes or when starting from Conf IV. This is confirmed in the top panel of Figure 19.9 by the much broader peaks obtained for Conf IV whose dynamics at harmonic zero point energy (even if less than 1 ps long) moves fast back and forth to Conf I. This aspect allows us to point out a couple of peculiar features of semiclassical dynamics. One is that the dynamics permits us to visit the effective portion of the potential energy surface and in this way construction of a global PES is not needed and ab initio on-the-fly calculations are fully legitimated. The second characteristic is that power spectra have a band shape determined by the dynamics. Results for Conf IV are corroborated by experiments performed at 13 K in argon matrices, which are much more difficult and less resolved for Conf IV [60]. The semiclassical calculations performed for glycine are based on an ab initio molecular dynamics at DFT-B3LYP level of theory with aVDZ basis set associated to MC SCIVR in full dimensionality. Results are in good agreement with experiments, and the interested reader can find all the details in Ref. [47]. Very recent work in the field has permitted calculation of the semiclassical IR spectrum of glycine [71].

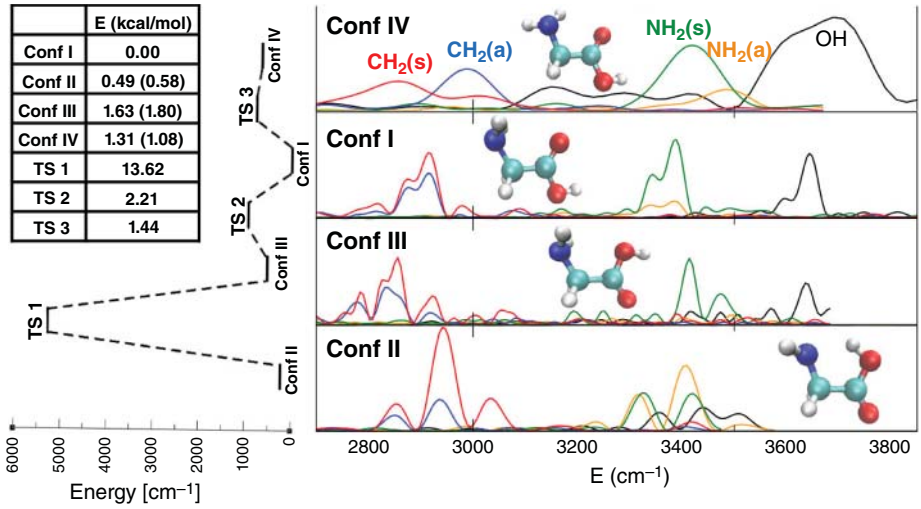


Figure 19.9 Left: Energetics of glycine calculated at DFT-B3LYP level of theory with aVDZ basis set. Energy values in parentheses include the harmonic zero-point energy contribution. Right: High frequency range vibrational power spectra for the four conformers of glycine. Peaks identify the OH (black), NH_2 asymmetric (orange), NH_2 symmetric (green), CH_2 asymmetric (blue), and CH_2 symmetric (red) stretches.

Another interesting application of semiclassical spectroscopy involving glycine is represented by the supra-molecular system made of protonated glycine tagged with hydrogen molecules. The goal is to reproduce and explain on a quantum mechanical basis the intriguing features of a recent experiment by Williams, Rizzo and coworkers [61]. In particular, in the experiment, the OH stretch signal presents twin peaks, separated by about 50 cm^{-1}, where it appears that at least three hydrogen molecules are present. The reason is that while the first pair of H_2 molecules interacts with the amino group of the protonated glycine, the third H_2 moiety can still interact either with the NH_3 group or with the hydroxyl. In this latter instance, the OH stretch is weakened and a second red-shifted peak is found in the experiment consistently with the mixture of the two conformers. Another feature of Williams and Rizzo's spectra is that the lowest in frequency among the three NH_3 bands is more and more blue-shifted as the number of tagging H_2 molecules increases, with a second peak appearing when three hydrogen molecules are involved. The blue shift is of about 60 cm^{-1}. This time the effect can be explained with a re-orientation of the NH_3 group which weakens the intramolecular hydrogen bond leading to a blue shift of the NH_3 signal. A theoretical description at the harmonic level fails even upon an ad-hoc scaling of frequencies calibrated on the OH stretch. In particular, the blue shift of the NH_3 peak is too overestimated predicting a complete cleavage of the intramolecular hydrogen bond instead of its simple weakening. This drawback is found also in a classical simulation based on the Fourier transform of the velocity–velocity correlation function. Conversely, a semiclassical description is able to gather both features, and frequencies are found within 20–30 cm^{-1} of the experiment. This SC simulation has been performed by means of the MC-DC-SCIVR technique interfaced to an ab initio on-the-fly molecular dynamics less than 1 ps long at DFT-B3LYP level of theory and with aVDZ basis set. Inclusion of dispersion corrections has been considered, but without any relevant gain in accuracy [62].

So far applications have been focused on power spectra that describe vibrational features on the ground electronic state, but semiclassical approaches can also be employed to simulate vibronic emission and absorption spectra that involve different electronic surfaces. A key example is given by the on-the-fly photoemission spectra of oligothiophenes calculated by the Vaníček group on the basis of the Fourier transform of a dipole–dipole autocorrelation function in Franck–Condon approximation (i.e., vertical electronic transition) for the initial wave packet [48]. The working formula is similar to Eq. (19.6) with the roles of the electronic ground (S_0) and excited (S_1) states interchanged, and a cubic dependence on the energy difference in front of the Fourier integration. A Thawed Gaussian approach was employed, and the emission spectrum was obtained upon short-time evolution ($\approx$ 200 fs) of the vibrational ground state of S_1 on the S_0 surface. This dynamical approach permits us to go beyond a static global harmonic approximation and to account for anharmonicities. Energy calculations and geometry optimization on the excited state surface were performed with TD-DFT, CAM-B3LYP functional and 6-31+G** basis set. The DFT level of theory was employed instead for the dynamics on the ground electronic state. Figure 19.10 reports the semiclassical photoemission spectra of a series of oligothiophenes and shows excellent agreement with the experimental results. Furthermore, in this work a strategy has been implemented to assess couplings between vibrational degrees of freedom during the dynamics, and their contribution to the spectral features. The same strategy permits us also to generate a partially uncoupled dynamics in reduced dimensionality, which may serve for application of more refined semiclassical or quantum techniques. The approach is different from DC SCIVR (which is based on a frozen instead of thawed Gaussian propagator), but it makes clear that DC SCIVR as well as the other techniques described in the chapter have the potential to also be applied successfully to the simulation of photoemission (absorption) spectra.

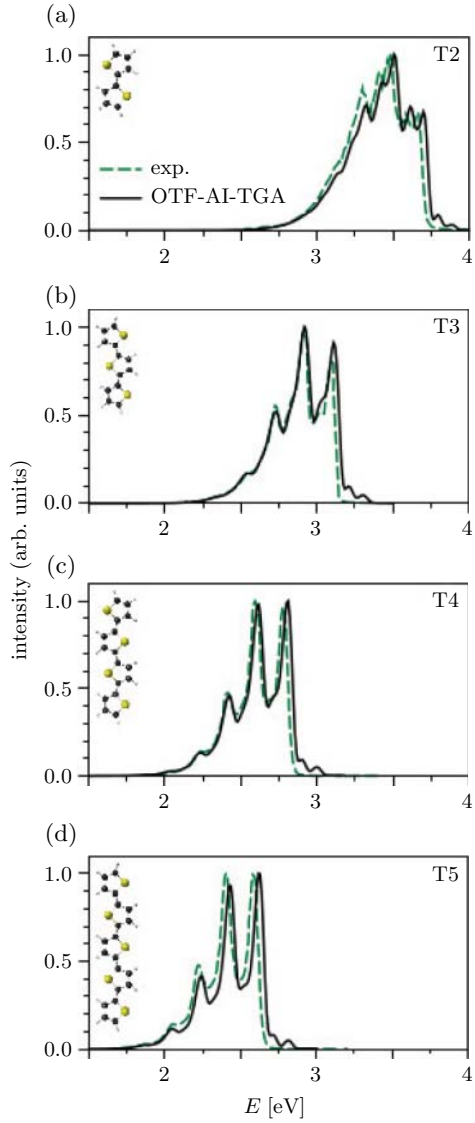


Figure 19.10 Photoemission spectra for oligothiophenes of increasing size. "OTF-AI-TGA" stands for on-the-fly ab initio thawed Gaussian approximation. Reproduced from Marius Wehrle, Miroslav Šulc, and Jiří Vaníček *The Journal of Chemical Physics* **140**, 244114 (2014), with the permission of AIP Publishing.

As a final application we consider another very recent work by the Vaníček group concerning the photoabsorption spectrum of benzene [63]. The relevant equation is Eq. (19.6), but in this case the Condon approximation cannot be invoked because the ($\tilde{A}^1\ B_{2u} \leftarrow \tilde{X}^1\ A_{1g}$) electronic transition is symmetry forbidden. However, it is vibronically allowed due to the contribution of the gradient of the transition dipole. In order to simulate the absorption spectrum of benzene one has to resort to the more refined Herzberg–Teller approximation in which the dipole depends linearly on the nuclear coordinates, i.e., $\vec{\mu}(\vec{R}) \approx \vec{\mu}(\vec{R}_{eq}) + \nabla_{\vec{R}}\vec{\mu}(\vec{R}_{eq}) \cdot (\vec{R} - \vec{R}_{eq})$. The Franck–Condon–Herzberg–Teller (FCHT) absorption spectrum of benzene (see Figure 19.11) was obtained by multiplying the initial Gaussian wave packet by a polynomial (linear) term. This

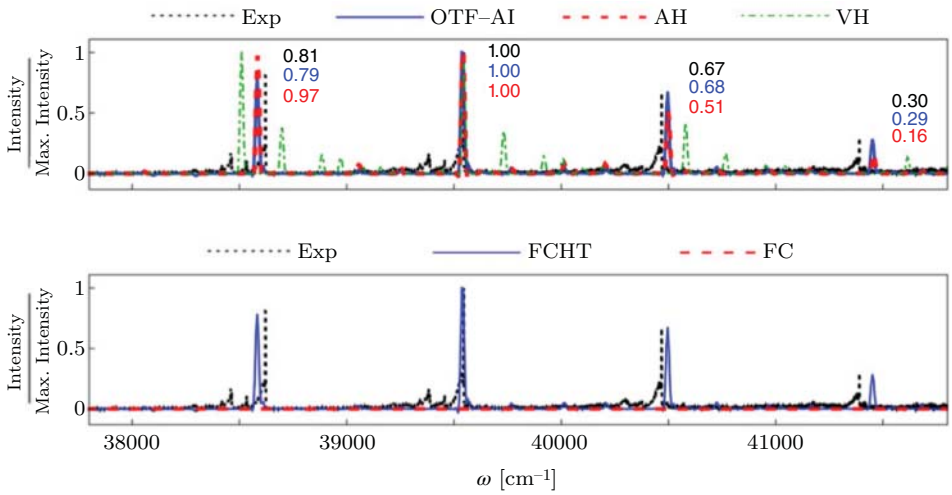


Figure 19.11 Semiclassical absorption spectra of benzene $\tilde{A}^1 B_{2u} \leftarrow \tilde{X}^1 A_{1g}$. In the top panel, a comparison between experiment and Franck–Condon–Herzberg–Teller (FCHT) spectra obtained from on-the-fly excited electronic state dynamics (OTF-AI), static adiabatic harmonic (AH), and static vertical harmonic (VH) approximations is presented. Scaled intensities are reported for each approach next to the peaks. In the bottom panel, the experiment is compared again to the FCHT spectrum from on-the-fly dynamics and to the Franck–Condon (FC) spectrum in Condon approximation. Note that the latter is zero because the electronic transition is symmetry forbidden. Reproduced from Aurélien Patoz, Tomislav Begušić, and Jiří Vaníček *The Journal of Physical Chemistry Letters* **9**, 2367 (2018). https://pubs.acs.org/doi/pdf/10.1021/acs.jpclett.8b00827. Requests for further reuse should be directed to the ACS.

permitted adoption of ab initio on-the-fly thawed Gaussian semiclassical dynamics with inclusion of the Herzberg–Teller approximation. The electronic excited state dynamics was performed for 2 ps with TD-DFT, B3LYP functional, and 6-31+G** basis set. DFT was employed for the optimization at the electronic ground state. The importance of a semiclassical approach able to account for the anharmonicities of the potential is pointed out in Figure 19.11, where a comparison between FCHT spectra adopting on-the-fly ab initio dynamics and the static adiabatic harmonic (AH) and vertical harmonic approximations (VH) is presented. In the AH approximation the upper electronic surface is obtained by means of a second-order expansion of the potential around the minimum, while the VH approximation is similar but based on the equilibrium configuration of the lower electronic state. The semiclassical approach outperforms the other two very clearly, with the VH approach returning very inaccurate results. Also in this application a different semiclassical approach based on a frozen Gaussian propagator and the techniques previously illustrated can be adopted.

19.11 Summary and Conclusions

In this chapter we have presented the derivation of the basic van Vleck and Heller–Herman–Kluk–Kay semiclassical propagators followed by a description of some recently developed SC techniques tailored for spectroscopy simulations. The multiple coherent states approach permits us to adopt ab initio on-the-fly molecular dynamics for semiclassical simulations with the advantage of avoiding construction of a global analytical potential energy surface. The divide-and-conquer technique allows us to investigate the quantum spectroscopic features of large molecular and supra-molecular

systems beating the so-called "curse of dimensionality". The simplified mixed semiclassical method is a promising tool that employs both frozen and thawed Gaussian semiclassical propagators to study vibrations of a molecule in a bath. Finally, some representative applications have been presented. They include power spectra of molecular and supra-molecular glycine (experimentally tagged with H_2 molecules), photoemission spectra of oligothiophenes, and the photoabsorption spectrum of benzene. In general, the described semiclassical techniques can be employed both for simulations on the electronic ground state and for calculations involving excited state dynamics.

To conclude we would like to review a few aspects of semiclassical spectroscopy that have either been recently solved or are the focus of current methodological developments. The bottleneck of semiclassical calculations lies in the determination of the Hessian matrix for the configurations visited along the trajectory, which is necessary for evolving the monodromy matrix elements and calculating the pre-exponential factor. This overhead has been alleviated by means of an interpolation scheme based on gradient estimates [64–66], and further improvements have recently led to an efficient strategy based on the dynamical construction of databases of Hessian matrices [67]. In the case of thawed Gaussian propagators, even an approximation based on a single Hessian has been recently proposed [68]. A potential issue of semiclassical simulations comes from the chaoticity of classical dynamics. Chaotic trajectories lead to a numerical loss of the unitarity of the determinant of the monodromy matrix and to unphysical values of the pre-exponential factor. This may endanger the entire calculation and often chaotic trajectories have to be discarded [20]. Some approximations to the pre-exponential factor have been proposed for overcoming this issue [69], and substantial progress on this issue has been achieved recently through development of an adiabatic switching technique for semiclassical spectroscopy [74]. As for the intensity of the calculated spectroscopic signals, apart from the possibility to simulate IR spectra starting from the dipole–dipole autocorrelation function, it is also possible to calculate them from the oscillator strengths once the wave functions are known. A procedure to determine the semiclassical wave functions expanding them on a harmonic basis set has been recently presented [70], followed by another approach capable of returning the semiclassical IR spectrum starting from power spectra even in the case of systems characterized by high densities of vibrational states [71].

Altogether we believe that semiclassical dynamics is a powerful tool for molecular spectroscopy with specific features that may make it the privileged approach for simulations of high dimensional systems. For instance, application of semiclassical spectroscopy has permitted the explaination of some features of two experiments which were left unanswered [62], while a very recent investigation focused on nucleobases [72]. Furthermore, the first calculations of vibrational semiclassical spectroscopy for molecules adsorbed on titania surfaces are now accessible [75]. We think that quantum effects should not be neglected a priori when investigating large systems spectroscopically. This is in part due to the quantum nature of spectroscopy itself and partly because of the not unusual presence of hydrogen bonds which require a quantum formalism to be described correctly. An advantage of semiclassical approaches is that they can be interfaced with ab initio molecular dynamics quite easily, since neither approximations to the potential energy nor ad hoc parameters are required. However, electronic calculations are computationally expensive and only a desirable speed up of electronic theory routines will permit employment of the highest levels of electronic theory in ab initio semiclassical simulations. Within the precision permitted by the electronic theory employed, an estimate of the accuracy of semiclassical simulations can be obtained by looking at the full width at half maximum of spectral peaks. This is generally found to be of the order of 20–25 cm^{-1} with occasional lower accuracy for particularly complex systems or spectral features, which may require further refinement. For example, a preliminary application of semiclassical spectroscopy to the Zundel cation $(H_5O_2)^+$ has demonstrated that accuracy is good even for this

Table 19.2 Principal strengths and weaknesses of semiclassical spectroscopy.

Strengths
+ Quantum effects from classical dynamics simulations.
+ Real potential energy, no approximations.
+ Interface to both analytical PESs and ab initio molecular dynamics.
+ No tunable nor ad hoc parameters.
+ Applicability to ground and excited electronic states.
+ Applicability to high dimensional systems.
Weaknesses
– Hessian calculation required.
– Classical chaotic trajectories may hamper simulations.
– Low ab initio level of electronic theory due to computational overhead.

floppy molecule [50]. A more refined and targeted study of this particular chemical species has been undertaken [76], leading to results of accuracy comparable to that of quantum calculations. A general formula to estimate the error with respect to the exact quantum mechanical result is not available, but formalisms to correct the semiclassical estimate towards the exact quantum mechanical result have been proposed by Kay and Pollak [9, 73].

Finally, we provide a brief summary of the main strengths and weaknesses of semiclassical spectroscopy, as reported in Table 19.2.

Acknowledgments

We thank past and current doctoral students and postdoctoral fellows of the Ceotto research group at Università degli Studi di Milano for their efforts and enthusiasm in developing semiclassical dynamics techniques and for revising this book chapter. In alphabetical order, they are: Dr. Chiara Aieta, Dr. Gianluca Bertaina, Dr. Max Buchholz, Dr. Marco Cazzaniga, Dr. Giovanni Di Liberto, Dr. Fabio Gabas, Dr. Huaqing Li, Dr. Ágnes Nóra Mahmoud, Dr. Marco Micciarelli, Alessandro Rognoni, and Dr. Jaime Suarez. Dr. Frank Grossmann is also warmly thanked for his careful revision. We thank Prof. Jiri Vaníček for useful discussions about the application Section. We acknowledge financial support from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant Agreement No. [647107]– SEMICOMPLEX– ERC-2014-CoG).

Bibliography

1 Grossmann, F. (2013). *Theoretical Femtosecond Physics: Atoms and Molecules in Strong Laser Fields*. Springer.

2 Qu, C., Yu, Q., and Bowman, J.M. (2018). Permutationally invariant potential energy surfaces. *Annu. Rev. Phys. Chem.* 69 (1): 151–175.

3 Heller, E.J. (1981). The semiclassical way to molecular spectroscopy. *Acc. Chem. Res.* 14 (12): 368–375.

4 Heller, E.J. (1978). Quantum corrections to classical photodissociation models. *J. Chem. Phys.* 68 (5): 2066–2075.

5 Miller, W.H. (2005). Quantum dynamics of complex molecular systems. *Proc. Natl. Acad. Sci. U. S. A.* 102 (19): 6660–6664.

6 Miller, W.H. (2001). The semiclassical initial value representation: a potentially practical way for adding quantum effects to classical molecular dynamics simulations. *J. Phys. Chem. A* 105 (13): 2942–2955.

7 Berry, M.V. and Mount, K.E. (1972). Semiclassical approximations in wave mechanics. *Rep. Prog. Phys.* 35 (1): 315.

8 Miller, W.H. (1974). Classical-limit quantum mechanics and the theory of molecular collisions. *Adv. Chem. Phys.* 25 (1): 69–177.

9 Kay, K.G. (2006). The Herman–Kluk approximation: derivation and semiclassical corrections. *Chem. Phys.* 322 (1–2): 3–12.

10 Feynman, R.P. and Ro Hibbs, A. (1965). *Quantum Mechanics and Path Integrals*. McGraw-Hill.

11 Swenson, D. (2011). *Quantum Effects from Classical Trajectories: New Methodologies and Applications for Semiclassical Dynamics Ab initio Molecular Dynamics: Basic Theory and Advanced Methods*. Ph.D. Dissertation, University of California, Berkeley.

12 Gutzwiller, M.C. (1967). Phase-integral approximation in momentum space and the bound states of an atom. *J. Math. Phys.* 8 (10): 1979–2000.

13 John, H. and Vleck, V. (1928). The correspondence principle in the statistical interpretation of quantum mechanics. *Proc. Natl. Acad. Sci.* 14 (2): 178–188.

14 Maslov, V.P. and Fedoriuk, M.V. (1981). *Semi-Classical Approximation in Quantum Mechanics*. Boston: Reidel.

15 Miller, W.H. (1970). Classical S matrix: numerical application to inelastic collisions. *J. Chem. Phys.* 53 (9): 3578–3587.

16 Heller, E.J. (1981). Frozen Gaussians: a very simple semiclassical approximation. *J. Chem. Phys.* 75 (6): 2923–2931.

17 Herman, M.F. and Kluk, E. (1984). A semiclassical justification for the use of non-spreading wave packets in dynamics calculations. *Chem. Phys.* 91 (1): 27–34.

18 Kay, K.G. (1994). Integral expressions for the semiclassical time-dependent propagator. *J. Chem. Phys.* 100 (6): 4377–4392.

19 Kay, K.G. (1994). Numerical study of semiclassical initial value methods for dynamics. *J. Chem. Phys.* 100 (6): 4432–4445.

20 Kay, K.G. (1994). Semiclassical propagation for multidimensional systems by an initial value method. *J. Chem. Phys.* 101 (3): 2250–2260.

21 Baranger, M., de Aguiar, M.A.M., Keck, F. et al. (2001). Semiclassical approximations in phase space with coherent states. *J. Phys. A* 34 (36): 7227.

22 Weissman, Y. (1982). Semiclassical approximation in the coherent states representation. *J. Chem. Phys.* 76 (8): 4067–4079.

23 Brewer, M.L., Hulme, J.S., and Manolopoulos, D.E. (1997). Semiclassical dynamics in up to 15 coupled vibrational degrees of freedom. *J. Chem. Phys.* 106 (12): 4832–4839.

24 Miller, W.H. (2002). An alternate derivation of the Herman Kluk (coherent state) semiclassical initial value representation of the time evolution operator. *Mol. Phys.* 100 (4): 397–400.

25 Filinov, V.S. (1986). Calculation of the Feynman integrals by means of the Monte Carlo method. *Nucl. Phys.* 271 (3): 717–725.

26 Makri, N. and Miller, W.H. (1988). Monte Carlo path integration for the real time propagator. *J. Chem. Phys.* 89 (4): 2170–2177.

27 Kluk, E., Herman, M.F., and Davis, H.L. (1986). Comparison of the propagation of semiclassical frozen Gaussian wave functions with quantum propagation for a highly excited anharmonic oscillator. *J. Chem. Phys.* 84 (1): 326–334.

28 Heller, E.J. (1991). Cellular dynamics: a new semiclassical approach to time-dependent quantum mechanics. *J. Chem. Phys.* 94 (4): 2723–2729.

29 Zambrano, E., Sulc, M., and Vaníček, J. (2013). Improving the accuracy and efficiency of time-resolved electronic spectra calculations: cellular dephasing representation with a prefactor. *J. Chem. Phys.* 139 (5): 054109.

30 Sulc, M. and Vaníček, J. (2012). Accelerating the calculation of time-resolved electronic spectra with the cellular dephasing representation. *Mol. Phys.* 110 (9-10): 945–955.

31 Wang, H., Manolopoulos, D.E., and Miller, W.H. (2001). Generalized Filinov transformation of the semiclassical initial value representation. *J. Chem. Phys.* 115 (14): 6317–6326.

32 Church, M.S., Antipov, S.V., and Ananth, N. (2017). Validating and implementing modified Filinov phase filtration in semiclassical dynamics. *J. Chem. Phys.* 146 (23): 234104.

33 Elran, Y. and Kay, K.G. (1999). Improving the efficiency of the Herman–Kluk propagator by time integration. *J. Chem. Phys.* 110 (8): 3653–3659.

34 Elran, Y. and Kay, K.G. (1999). Time-integrated form of the semiclassical initial value method. *J. Chem. Phys.* 110 (18): 8912–8918.

35 Kaledin, A.L. and Miller, W.H. (2003). Time averaging the semiclassical initial value representation for the calculation of vibrational energy levels. *J. Chem. Phys.* 118 (16): 7174–7182.

36 Kaledin, A.L. and Miller, W.H. (2003). Time averaging the semiclassical initial value representation for the calculation of vibrational energy levels. II. Application to H_2,CO,NH_3,CH_4,CH_2D_2. *J. Chem. Phys.* 119 (6): 3078–3084.

37 Tamascelli, D., Dambrosio, F.S., Conte, R., and Ceotto, M. (2014). Graphics processing units accelerated semiclassical initial value representation molecular dynamics. *J. Chem. Phys.* 140 (17): 174109.

38 Pratihar, S., Ma, X., Homayoon, Z. et al. (2017). Direct chemical dynamics simulations. *J. Am. Chem. Soc.* 139 (10): 3570–3590.

39 Ceotto, M., Atahan, S., Tantardini, G.F., and Aspuru-Guzik, A. (2009). Multiple coherent states for first-principles semiclassical initial value representation molecular dynamics. *J. Chem. Phys.* 130 (23): 234113.

40 De Leon, N. and Heller, E.J. (1983). Semiclassical quantization and extraction of eigenfunctions using arbitrary trajectories. *J. Chem. Phys.* 78: 4005–4017.

41 Ceotto, M., Tantardini, G.F., and Aspuru-Guzik, A. (2011). Fighting the curse of dimensionality in first-principles semiclassical calculations: non-local reference states for large number of dimensions. *J. Chem. Phys.* 135 (21): 214108.

42 Ceotto, M., Dell' Angelo, D., and Tantardini, G.F. (2010). Multiple coherent states semiclassical initial value representation spectra calculations of lateral interactions for CO on Cu (100). *J. Chem. Phys.* 133 (5): 054701.

43 Ceotto, M., Valleau, S., Tantardini, G.F., and Aspuru-Guzik, A. (2011). First principles semiclassical calculations of vibrational eigenfunctions. *J. Chem. Phys.* 134 (23): 234103.

44 Ceotto, M., Atahan, S., Shim, S. et al. (2009). First-principles semiclassical initial value representation molecular dynamics. *Phys. Chem. Chem. Phys.* 11: 3861–3867.

45 Dressler, S. and Thiel, W. (1997). Anharmonic force fields from density functional theory. *Chem. Phys. Lett.* 273 (1): 71–78.

46 Conte, R., Aspuru-Guzik, A., and Ceotto, M. (2013). Reproducing deep tunneling splittings, resonances, and quantum frequencies in vibrational spectra from a handful of direct ab initio semiclassical trajectories. *J. Phys. Chem. Lett.* 4 (20): 3407–3412.

47 Gabas, F., Conte, R., and Ceotto, M. (2017). On-the-fly ab initio semiclassical calculation of glycine vibrational spectrum. *J. Chem. Theory Comput.* 13: 2378.

48 Wehrle, M., Sulc, M., and Vanicek, J. (2014). On-the-fly ab initio semiclassical dynamics: identifying degrees of freedom essential for emission spectra of oligothiophenes. *J. Chem. Phys.* 140 (24): 244114.

49 Ceotto, M., Di Liberto, G., and Conte, R. (2017). Semiclassical "Divide-and-Conquer" method for spectroscopic calculations of high dimensional molecular systems. *Phys. Rev. Lett.* 119 (1): 010401.

50 Di Liberto, G., Conte, R., and Ceotto, M. (2018). "divide and conquer" semiclassical molecular dynamics: a practical method for spectroscopic calculations of high dimensional molecular systems. *J. Chem. Phys.* 148 (1): 014307.

51 Di Liberto, G., Conte, R., and Ceotto, M. (2018). "divide-and-conquer" semiclassical molecular dynamics: an application to water clusters. *J. Chem. Phys.* 148 (10): 104302.

52 Heller, E.J. (1975). Time dependent approach to semiclassical dynamics. *J. Chem. Phys.* 62 (4): 1544–1555.

53 Buchholz, M., Grossmann, F., and Ceotto, M. (2016). Mixed semiclassical initial value representation time-averaging propagator for spectroscopic calculations. *J. Chem. Phys.* 144: 094102.

54 Grossmann, F. (2006). A semiclassical hybrid approach to many particle quantum dynamics. *J. Chem. Phys.* 125 (1).

55 Buchholz, M., Grossmann, F., and Ceotto, M. (2017). Application of the mixed time-averaging semiclassical initial value representation method to complex molecular spectra. *J. Chem. Phys.* 147 (16): 164110.

56 Buchholz, M., Grossmann, F., and Ceotto, M. (2018). Simplified approach to the mixed time-averaging semiclassical initial value representation for the calculation of dense vibrational spectra. *J. Chem. Phys.* 148 (11): 114107.

57 Valiev, M., Bylaska, E.J., Govind, N. et al. (2010). NWChem: a comprehensive and scalable open-source solution for large scale molecular simulations. *Comput. Phys. Commun.* 181 (9): 1477–1489.

58 Lourderaj, U., Sun, R., Kohale, S.C. et al. (2014). The venus/NWChem software package. Tight coupling between chemical dynamics simulations and electronic structure theory. *Comput. Phys. Commun.* 185 (3): 1074–1080.

59 Ma, X., Di Liberto, G., Conte, R. et al. (2018). A quantum mechanical insight into SN2 reactions: semiclassical initial value representation calculations of vibrational features of the $Cl–CH_3Cl$ pre-reaction complex with the venus suite of codes. *J. Chem. Phys.* 149 (16): 164113.

60 Stepanian, S.G., Reva, I.D., Radchenko, E.D. et al. (1998). Matrix-isolation infrared and theoretical studies of the glycine conformers. *J. Phys. Chem. A* 102 (6): 1041–1054.

61 Masson, A., Williams, E.R., and Rizzo, T.R. (2015). Molecular hydrogen messengers can lead to structural infidelity: a cautionary tale of protonated glycine. *J. Chem. Phys.* 143 (10): 104313.

62 Gabas, F., Di Liberto, G., Conte, R., and Ceotto, M. (2018). Protonated glycine supramolecular systems: the need for quantum dynamics. *Chem. Sci.* 9: 7894–7901.

63 Patoz, A., Begusic, T., and Vanicek, J. (2018). On-the-fly ab initio semiclassical evaluation of absorption spectra of polyatomic molecules beyond the condon approximation. *J. Phys. Chem. Lett.* 9 (9): 2367–2372.

64 Wu, H., Rahman, M., Wang, J. et al. (2010). Higher-accuracy schemes for approximating the hessian from electronic structure calculations in chemical dynamics simulations. *J. Chem. Phys.* 133 (7): 074101.

65 Zhuang, Y., Siebert, M.R., Hase, W.L. et al. (2012). Evaluating the accuracy of Hessian approximations for direct dynamics simulations. *J. Chem. Theory Comput.* 9 (1): 54–64.

66 Ceotto, M., Zhuang, Y., and Hase, W.L. (2013). Accelerated direct semiclassical molecular dynamics using a compact finite difference Hessian scheme. *J. Chem. Phys.* 138 (5): 054116.

67 Conte, R., Gabas, F., Botti, G. et al. (2019). Semiclassical vibrational spectroscopy with hessian databases. *J. Chem. Phys.* 150: 244118.

68 Begusic, T., Cordova, M., and Vanicek, J. (2019). Single-Hessian thawed Gaussian approximation. *J. Chem. Phys.* 150 (15): 154117.

69 Di Liberto, G. and Ceotto, M. (2016). The importance of the pre-exponential factor in semiclassical molecular dynamics. *J. Chem. Phys.* 145: 144107.

70 Micciarelli, M., Conte, R., Suarez, J., and Ceotto, M. (2018). Anharmonic vibrational eigenfunctions and infrared spectra from semiclassical molecular dynamics. *J. Chem. Phys.* 149: 064115.

71 Micciarelli, M., Gabas, F., Conte, R., and Ceotto, M. (2019). An effective semiclassical approach to ir spectroscopy. *J. Chem. Phys.* 150 (18): 184113.

72 Gabas, F., Di Liberto, G., and Ceotto, M. (2019). Vibrational investigation of nucleobases by means of divide-and-conquer semiclassical dynamics. *J. Chem. Phys.* 150: 224107.

73 Pollak, E. (2007). *The Semiclassical Initial Value Series Representation of the Quantum Propagator*, 259–271. Berlin, Heidelberg: Springer Berlin Heidelberg.

74 Conte, R., Parma, L., Aieta, C. et al. (2019). Improved semiclassical dynamics through adiabatic switching trajectory sampling. *J. Chem. Phys.* 151: 214107.

75 Cazzaniga, M., Micciarelli, M., Moriggi, F. et al. (2020). Anharmonic calculations of vibrational spectra for molecular adsorbates: a divide-and-conquer semiclassical molecular dynamics approach. *J. Chem. Phys.* 152: 104104.

76 Bertaina, G., Di Liberto, G., and Ceotto, M. (2019). Reduced rovibrational coupling Cartesian dynamics for semiclassical calculations: application to the spectrum of the Zundel cation. *J. Chem. Phys.* 151: 114307.

20

Path-Integral Approaches to Non-Adiabatic Dynamics

Maximilian A. C. Saller, Johan E. Runeson, and Jeremy O. Richardson

Laboratory of Physical Chemistry, ETH Zurich, 8093 Zurich, Switzerland

Abstract

In this chapter, we describe methods for simulating non-adiabatic dynamics based on the path-integral formulation of quantum mechanics. In order to employ trajectory calculations to a system of more than one electronic state, we introduce the mapping formalism and explain how this approach can be used with linearized semiclassical or ring-polymer molecular dynamics.

20.1 Introduction

In order to simulate non-adiabatic processes in complex systems of many degrees of freedom, a fully quantum mechanical approach is impossible to apply in practice. Instead, we would like to make a compromise between accuracy and efficiency and in particular make use of extensions of the molecular dynamics technique, which has proved itself invaluable for the simulation of adiabatic systems. The path-integral approach to quantum mechanics [1] leads to a formally exact description of non-adiabatic dynamics from which it is possible to make approximations leading to practical simulation methods. In this chapter, we discuss a class of methodologies which utilize molecular dynamics trajectories but include the quantum mechanical effects of nuclear delocalization and tunneling as well as allowing for non-adiabatic transitions. This is done using the non-adiabatic mapping formalism in conjunction with path-integral molecular dynamics.

Here we are interested in simulating the quantum dynamics of the following nuclear Hamiltonian, written in the diabatic representation (see Chapter 1),

$$\hat{H} = \frac{\hat{\vec{P}}^2}{2M} + \hat{U} + \hat{V}, \tag{20.1}$$

where M is the nuclear mass, the state-independent potential is $\hat{U} = U(\hat{\vec{R}})$ and the state-dependent potential is

$$\hat{V} = V_1(\hat{\vec{R}})|1\rangle\langle 1| + V_2(\hat{\vec{R}})|2\rangle\langle 2| + \Delta(\hat{\vec{R}})(|1\rangle\langle 2| + |2\rangle\langle 1|), \tag{20.2}$$

where $|1\rangle$ and $|2\rangle$ are the electronic states, and $\vec{R}$ and $\vec{P}$ are the Cartesian coordinate and conjugate momentum vectors of length F, corresponding to the nuclear degrees of freedom.

Quantum Chemistry and Dynamics of Excited States: Methods and Applications,
First Edition. Edited by Leticia González and Roland Lindh.

The state-dependent potential can also be written in matrix form as

$$\mathrm{V}(\vec{R}) = \begin{pmatrix} V_1(\vec{R}) & \Delta(\vec{R}) \\ \Delta(\vec{R}) & V_2(\vec{R}) \end{pmatrix}, \tag{20.3}$$

where the basis functions are the electronic states. This two-state system can easily be extended to describe problems of more than two electronic states, as can the methodologies presented below for simulating the dynamics.

The two diabatic potentials are thus $U(\vec{R}) + V_1(\vec{R})$ and $U(\vec{R}) + V_2(\vec{R})$ and are coupled together by $\Delta(\vec{R})$. While this separation into $\hat{U}$ and $\hat{V}$ is not unique, the quantum dynamics are independent of this choice. The dynamics are also independent of the representation, such that it would be possible to rotate the electronic basis to the adiabatic representation, thus introducing non-adiabatic coupling into the kinetic energy, but resulting in equivalent dynamics.

To simplify the presentation, we assume that the coordinates have been mass scaled such that the same mass, M, can be used for all degrees of freedom.

When studying complex systems, especially those in the condensed phase, the observables of interest are typically correlation functions defined by

$$C_{AB}(t) = \langle \hat{A}(0)\hat{B}(t)\rangle_{\hat{\rho}} = \mathrm{Tr}[\hat{\rho}\hat{A}\mathrm{e}^{\mathrm{i}\hat{H}t/\hbar}\hat{B}\mathrm{e}^{-\mathrm{i}\hat{H}t/\hbar}]. \tag{20.4}$$

Here the dynamics start from an initially-prepared state defined by the density matrix, $\hat{\rho}$. In general this may be a non-equilibrium state, but a special class of correlation function is initialized from the thermal equilibrium state, $\rho_{\mathrm{eq}} = \mathrm{e}^{-\beta\hat{H}}/Z$, where the partition function is $Z = \mathrm{Tr}[\mathrm{e}^{-\beta\hat{H}}]$ and $\beta = 1/kT$ is the inverse temperature. This defines the equilibrium correlation function

$$C^{\mathrm{eq}}_{AB}(t) = \langle \hat{A}(0)\hat{B}(t)\rangle_{\mathrm{eq}} = \frac{1}{Z}\,\mathrm{Tr}[\mathrm{e}^{-\beta\hat{H}}\hat{A}\mathrm{e}^{\mathrm{i}\hat{H}t/\hbar}\hat{B}\mathrm{e}^{-\mathrm{i}\hat{H}t/\hbar}]. \tag{20.5}$$

The same information content is also contained in other representations, such as the Kubo-transformed correlation function [2],

$$\tilde{C}_{AB}(t) = \frac{1}{Z\beta}\int_0^\beta \mathrm{Tr}[\mathrm{e}^{-(\beta-\lambda)\hat{H}}\hat{A}\mathrm{e}^{-\lambda\hat{H}}\mathrm{e}^{\mathrm{i}\hat{H}t/\hbar}\hat{B}\mathrm{e}^{-\mathrm{i}\hat{H}t/\hbar}]\,\mathrm{d}\lambda, \tag{20.6}$$

which has the particular advantage of being real-valued at all times and is the natural form which leads to the approximation of ring-polymer molecular dynamics (RPMD) [3] as will be explained below. This transform can be thought of as an average over every possible splitting of the $\mathrm{e}^{-\beta\hat{H}}$ operator, with λ acting as the arbitrary integration variable. Once the Kubo-transformed correlation function is known, it is easy to obtain the ordinary equilibrium correlation function, $C^{\mathrm{eq}}_{AB}(t)$, by scaling the former's Fourier transform [4],

$$\int C^{\mathrm{eq}}_{AB}\mathrm{e}^{-\mathrm{i}\omega t}\,\mathrm{d}t = \frac{\beta\hbar\omega}{1-\mathrm{e}^{-\beta\hbar\omega}}\int \tilde{C}_{AB}(t)\mathrm{e}^{-\mathrm{i}\omega t}\,\mathrm{d}t. \tag{20.7}$$

Many important experimental observables can be related to thermal correlation functions [5]. In particular, by taking a Fourier transform of the dipole–dipole correlation functions, we can obtain the vibronic spectrum [6, 7, 9, 10] and by integrating over the flux–flux correlation function, the rate of a non-adiabatic process can be found [11–17].

In some cases, a simpler time-dependent observable is of interest, which can be defined in the same way by simply setting the operator $\hat{A}$ to the identity. The expectation value of an operator $\hat{B}$ at time t later is then given by

$$P_B(t) = \langle \hat{B}(t)\rangle_{\hat{\rho}} = \mathrm{Tr}[\hat{\rho}\ \mathrm{e}^{\mathrm{i}\hat{H}t/\hbar}\hat{B}\mathrm{e}^{-\mathrm{i}\hat{H}t/\hbar}]. \tag{20.8}$$

Because the equilibrium distribution commutes with the time propagators, the equilibrium time average of an observable is constant, $\langle \hat{B}(t) \rangle_{\mathrm{eq}} = \langle \hat{B} \rangle_{\mathrm{eq}}$. Therefore, this observable is typically only studied for non-equilibrium initial conditions.

In this chapter, we describe how these time-dependent observables and correlation functions can be approximately simulated in an efficient way. The major difficulty for simulating the dynamics of this system is due to the fact that the electronic states are described by a discrete basis, $|1\rangle$ or $|2\rangle$, whereas the nuclear degrees of freedom have a continuum of position and momentum states. Typically an approximation is made to treat the nuclei classically and the electronic states quantum mechanically [18]. However, we would like to treat the nuclear and electronic degrees of freedom on an equal footing, in order to also include nuclear quantum effects. We will, therefore, introduce the mapping approach [19, 20] to rigorously convert the discrete electronic basis into a continuous basis. This will then be used first to obtain linearized path-integral dynamics and then ring-polymer molecular dynamics. We show results of non-equilibrium and equilibrium correlation functions applied to example systems and discuss their behaviour.

20.2 Semiclassical Theory

There have been a number of theories developed for simulating semiclassical non-adiabatic dynamics [21], including surface hopping approaches [22] and mean-field dynamics [23]. Here we will introduce the mapping formalism for representing discrete electronic states by continuous phase-space variables. This will allow us to obtain classical equations of motion for these variables and to calculate the correlation functions as averages over ensembles of the trajectories.

20.2.1 Mapping Approach

A classical model of non-adiabatic transitions was created by Meyer and Miller [19] in terms of an additional set of variables to describe the electronic state. This model was later formalized by Stock and Thoss [20] and shown to be rigorously derived from quantum mechanics. The proof is based on the equivalence between an L-level system and a set of L harmonic-oscillator creation and annihilation operators, $\hat{a}_\nu^\dagger$ and $\hat{a}_\nu$. The operators and states are mapped to the new representation according to the following relations:

$$|\nu\rangle\langle\mu| \mapsto \hat{a}_\nu^\dagger \hat{a}_\mu \tag{20.9a}$$

$$|\nu\rangle \mapsto |0_1, \cdots, 1_\nu, \cdots, 0_L\rangle. \tag{20.9b}$$

The original electronic states are thus represented by singly-excited harmonic oscillator states in L degrees of freedom, and the operators act on these states by moving the excitation between different degrees of freedom. One can show that this mapping preserves the commutation relation of the original system,

$$[|\nu\rangle\langle\mu|, |\nu'\rangle\langle\mu'|] = |\nu\rangle\langle\mu'|\delta_{\mu\nu'} - |\nu'\rangle\langle\mu|\delta_{\mu'\nu}. \tag{20.10}$$

The advantage of changing the representation to one based on creation and annihilation operators of harmonic oscillators is that these can themselves be represented using position and momentum operators, $\hat{x}_\nu = (\hat{a}_\nu^\dagger + \hat{a}_\nu)/\sqrt{2}$ and $\hat{p}_\nu = \mathrm{i}(\hat{a}_\nu^\dagger - \hat{a}_\nu)/\sqrt{2}$, given in the continuous (dimensionless) position and momentum representations by

$$\langle \mathsf{x}|\nu\rangle = \frac{\sqrt{2}}{\pi^{L/4}}\, x_\nu\, \mathrm{e}^{-\frac{1}{2}|\mathsf{x}|^2}, \qquad \langle \mathsf{p}|\nu\rangle = \frac{-\mathrm{i}\sqrt{2}}{\pi^{L/4}}\, p_\nu\, \mathrm{e}^{-\frac{1}{2}|\mathsf{p}|^2}, \tag{20.11}$$

where $\mathsf{x} = (x_1, \cdots, x_L)$ and $\mathsf{p} = (p_1, \cdots, p_L)$ are vectors of the mapping positions and momenta. This is visualized in Figure 20.1.

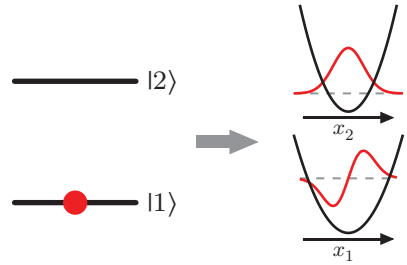


Figure 20.1 A schematic of the mapping approach. The two-state quantum system in state $|1\rangle$, as shown on the left, is mapped to a set of two harmonic oscillators, with an excitation in the mode corresponding to the populated state.

Using the fact that the commutator $[\hat{x}_\nu, \hat{p}_\mu] = \mathrm{i}\delta_{\nu\mu}$, the state operators are mapped to

$$|\nu\rangle\langle\nu| = \frac{1}{2}(\hat{x}_\nu^2 + \hat{p}_\nu^2 - 1) \tag{20.12a}$$

$$|\nu\rangle\langle\mu| + |\mu\rangle\langle\nu| = \hat{x}_\nu\hat{x}_\mu + \hat{p}_\nu\hat{p}_\mu \qquad \nu \neq \mu\,, \tag{20.12b}$$

and therefore the mapping Hamiltonian can be defined as

$$\hat{\mathcal{H}} = \frac{\hat{\vec{P}}^2}{2M} + \hat{U} + \frac{1}{2}[\hat{\mathrm{x}}^{\mathrm{T}}\hat{\mathrm{V}}\hat{\mathrm{x}} + \hat{\mathrm{p}}^{\mathrm{T}}\hat{\mathrm{V}}\hat{\mathrm{p}} - \mathrm{tr}\hat{\mathrm{V}}]. \tag{20.13}$$

The propagator based on this mapping Hamiltonian leads to equivalent dynamics to the unmapped Hamiltonian as $\langle\nu|\mathrm{e}^{-\mathrm{i}\hat{H}t/\hbar}|\mu\rangle = \langle 0_1, \cdots, 1_\nu, \cdots, 0_L|\mathrm{e}^{-\mathrm{i}\hat{\mathcal{H}}t/\hbar}|0_1, \cdots, 1_\mu, \cdots, 0_L\rangle$. The advantage of the mapping representation is that the Hamiltonian has a well-defined classical analogue, in which x_ν and p_ν are used not as operators, but as canonical variables that evolve according to Hamilton's equations.

In principle, the mapping presented above is formally exact. However, this is no longer true once classical equations of motions are employed. In particular, the true quantum dynamics are constrained to a physical subspace of singly-excited oscillators, although due to zero-point energy leakage, the classical dynamics may leave this subspace [24, 25]. Other mapping approaches have also been developed including using a spin degree of freedom rather than singly-excited harmonic oscillators. This gives rise to the same equations of motion but may have certain advantages in that the classical dynamics cannot leave the physical subspace [26].

While these definitions were given in the diabatic representation, there is an equivalent formulation of the mapping approach in the adiabatic representation. This is easily obtained using a canonical transformation and thus gives identical results [27, 28]. A middle ground is provided by quasi-diabatic representations which may lead to simpler methods when combined with on-the-fly ab initio calculation of the potential energy surfaces [29].

20.2.2 Linearized Semiclassical Dynamics

In order to obtain a practical simulation method, some approximations must be made to the quantum dynamics. In particular, we would like to employ a classical trajectory method to propagate the system defined in the mapping representation. To make a connection with the quantum correlation functions defined in Eq. (20.4), the quantum operators are first Wigner transformed into classical phase-space representations.

The Wigner transform of an operator $\hat{\mathcal{O}}$ is defined as

$$\mathcal{O}^{\mathrm{w}}(\vec{R}, \vec{P}, \mathrm{x}, \mathrm{p}) = \iint \mathrm{e}^{\mathrm{i}\vec{P}\cdot\vec{R}'/\hbar + \mathrm{i}\mathrm{p}\cdot\mathrm{x}'} \left\langle \vec{R} - \frac{\vec{R}'}{2}, \mathrm{x} - \frac{\mathrm{x}'}{2} \middle| \hat{\mathcal{O}} \middle| \vec{R} + \frac{\vec{R}'}{2}, \mathrm{x} + \frac{\mathrm{x}'}{2} \right\rangle \mathrm{d}\mathrm{x}'\ \mathrm{d}\vec{R}'\,, \tag{20.14}$$

where x' and $\vec{R}'$ are integration variables in the same space as $\{\mathsf{x}, \mathsf{p}\}$ and $\{\vec{R}, \vec{P}\}$ respectively, facilitating the Wigner transform over $\hat{\mathcal{O}}$. The exact quantum correlation function can be written in these terms as

$$C_{AB}(t) = \frac{1}{(2\pi\hbar)^F (2\pi)^L} \iiiint \rho^{\mathrm{W}}(\vec{R}, \vec{P}, \mathsf{x}, \mathsf{p}) A^{\mathrm{W}}(\vec{R}, \vec{P}, \mathsf{x}, \mathsf{p}) [\hat{B}(t)]^{\mathrm{W}}(\vec{R}, \vec{P}, \mathsf{x}, \mathsf{p}) \, \mathrm{d}\mathsf{x} \, \mathrm{d}\mathsf{p} \, \mathrm{d}\vec{R} \, \mathrm{d}\vec{P}. \tag{20.15}$$

So far, no approximation has been made, but even though the formula is written in terms of classical phase-space distributions, it is not a practical method because the Heisenberg representation of the quantum operator $\hat{B}(t)$ is only known at short times and is impractical to compute at long times. A simple approximation is obtained by using time-dependent classical trajectories to obtain the Wigner-transformed operator at later times. This results in the correlation function defined as

$$C_{AB}(t) \approx \frac{1}{(2\pi\hbar)^F (2\pi)^L} \iiiint \rho^{\mathrm{W}}(\vec{R}, \vec{P}, \mathsf{x}, \mathsf{p}) A^{\mathrm{W}}(\vec{R}, \vec{P}, \mathsf{x}, \mathsf{p}) B^{\mathrm{W}}(\vec{R}(t), \vec{P}(t), \mathsf{x}(t), \mathsf{p}(t)) \mathrm{d}\mathsf{x} \, \mathrm{d}\mathsf{p} \mathrm{d}\vec{R} \mathrm{d}\vec{P}, \tag{20.16}$$

where the time-dependent quantities are found by propagating the initial conditions according to the classical mapping Hamiltonian

$$\mathcal{H} = \frac{\vec{P}^2}{2M} + U(\vec{R}) + \frac{1}{2}[\mathsf{x}^{\mathrm{T}} \mathsf{V} \mathsf{x} + \mathsf{p}^{\mathrm{T}} \mathsf{V} \mathsf{p} - \mathrm{tr}\mathsf{V}]. \tag{20.17}$$

The equations of motion of the mapping variables are

$$\dot{\mathsf{x}} = \mathsf{V}\mathsf{p} \tag{20.18a}$$

$$\dot{\mathsf{p}} = -\mathsf{V}\mathsf{x} \,, \tag{20.18b}$$

and of the nuclear coordinates and momenta,

$$\dot{\vec{R}} = \frac{\vec{P}}{M} \tag{20.18c}$$

$$\dot{\vec{P}} = -\frac{\partial U}{\partial \vec{R}} - \frac{1}{2}\left[\mathsf{x}^{\mathrm{T}} \frac{\partial \mathsf{V}}{\partial \vec{R}} \mathsf{x} + \mathsf{p}^{\mathrm{T}} \frac{\partial \mathsf{V}}{\partial \vec{R}} \mathsf{p} - \mathrm{tr}\, \frac{\partial \mathsf{V}}{\partial \vec{R}}\right]. \tag{20.18d}$$

One important aspect to note about these equations is that $|\mathsf{x}|^2 + |\mathsf{p}|^2$ is a conserved quantity. The equations of motion can be solved iteratively using extensions of standard molecular dynamics [25]. An example calculation performed by this method will be presented in the following section.

This approximate approach can be derived rigorously in the semiclassical limit by linearizing the path-integral representation of the operator $\hat{B}(t)$. This approach is, therefore, known as the linearized semiclassical initial-value representation (LSC-IVR) [30–32]. It is also strongly related to the quantum-classical Liouville equation [33–35].

There also exist less drastic approximations where the path-integral is not fully linearized. These give semiclassical dynamics including phase contributions from the trajectories [20, 27, 36] as well as partially-linearized dynamics [15, 37–39]. Alternative approaches based on the same dynamics have also been suggested using binned conditions for the mapping variables [40].

20.3 Non-Equilibrium Dynamics

One of the key challenges in condensed matter physics is the description of the dynamics in dissipative quantum systems. The physical phenomena affected by the complexity of dissipation in

condensed phase systems include electron transfer, hydrogen tunneling as well as a number of others [41–43].

20.3.1 Spin-Boson Systems

The Spin-Boson model captures the key aspects needed to describe quantum dissipation. It has been widely studied to this end [13, 31, 44–51] owing to the relative simplicity of its Hamiltonian. The model consists of a two-level subsystem, coupled linearly to a bath, approximated as a set of harmonic oscillators.

The general form of the Spin-Boson Hamiltonian is written as a sum of subsystem, system-bath and bath Hamiltonians, $\hat{H}_s$, $\hat{H}_{sb}$ and $\hat{H}_b$ respectively.

$$\hat{H} = \underbrace{\varepsilon\hat{\sigma}_z + \Delta\hat{\sigma}_x}_{\hat{H}_s} + \underbrace{\sum_{j=1}^{F} c_j R_j \hat{\sigma}_z}_{\hat{H}_{sb}} + \underbrace{\sum_{j=1}^{F}\left[\frac{P_j^2}{2M_j} + \frac{1}{2}M_j\omega_j^2 R_j^2\right]\hat{I}}_{\hat{H}_b}, \tag{20.19}$$

where ε is the energy bias, Δ is the inter-state or electronic coupling and c_j are the system-bath, intra-state or vibronic coupling coefficients. F is the total number of bath modes, R_j and P_j denote the position and momentum of bath degree of freedom j, while M_j is its mass and ω_j its frequency.

Systems featuring two discrete states, of which the Spin-Boson Hamiltonian is a key example, are generally described by the Pauli matrices, which are

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \hat{\sigma}_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} \qquad \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{20.20}$$

and the identity $\hat{I}$. While convenient as descriptions of a two state system in their matrix form, in order to express the operators associated with the Pauli matrices in the mapping formalism, the equivalent expressions in Dirac notation are used,

$$\hat{\sigma}_x = |2\rangle\langle 1| + |1\rangle\langle 2| \qquad \hat{\sigma}_y = \mathrm{i}(|2\rangle\langle 1| - |1\rangle\langle 2|) \qquad \hat{\sigma}_z = |1\rangle\langle 1| - |2\rangle\langle 2|\,. \tag{20.21}$$

Given these definitions and the relation between state operators and creation and annihilation operators, given in Eq. (20.9a), the mapping representations of the Pauli matrix operators are

$$\hat{\sigma}_x \mapsto \hat{x}_1\hat{x}_2 + \hat{p}_1\hat{p}_2 \tag{20.22a}$$

$$\hat{\sigma}_y \mapsto \hat{x}_1\hat{p}_2 - \hat{p}_1\hat{x}_2 \tag{20.22b}$$

$$\hat{\sigma}_z \mapsto \frac{1}{2}(\hat{x}_1^2 + \hat{p}_1^2 - \hat{x}_2^2 - \hat{p}_2^2)\,. \tag{20.22c}$$

As mentioned in section Section 20.2.1, classical analogues to these operators can be arrived at by replacing the $\hat{\mathsf{x}}$ and $\hat{\mathsf{p}}$ operators with the classical phase-space variables, x and p, leading to a classical set of Pauli matrix representations, similar to Eqs. (20.22), such that for example, $\sigma_x^{cl}(\mathsf{x},\mathsf{p}) = x_1x_2 + p_1p_2$.

Another strategy for arriving at classical phase-space representations of quantum operators, already mentioned in section Section 20.2.2, utilizes the Wigner transform, defined as in Eq. (20.14). When applying this to the Spin-Boson Hamiltonian and therefore taking the transform over the Pauli matrix operators, a crucial distinction must be made which representation of these operators is used [25].

The Wigner transform of the Pauli operators, written in terms of the mapping variable operators, $\hat{\mathsf{x}}$ and $\hat{\mathsf{p}}$, shown in Eqs. (20.22), results in phase-space representations identical to the classical

equivalents of Eqs. (20.22), *e.g.*, $\sigma_x^{\mathrm{cl}}(\mathsf{x},\mathsf{p})$. If we instead use the Pauli operators in their state vector form, as shown in Eq. (20.21), that is projected onto the harmonic oscillator subspace, this results in a different set of phase-space representations:

$$\sigma_\alpha^{\mathrm{w}}(\mathsf{x},\mathsf{p}) = 16\ \tilde{\sigma}_\alpha(\mathsf{x},\mathsf{p})\ \mathrm{e}^{-(x_1^2+p_1^2+x_2^2+p_2^2)} \qquad \alpha \in \{x,y,z\}\ , \tag{20.23}$$

where

$$\tilde{\sigma}_x(\mathsf{x},\mathsf{p}) = x_1x_2 + p_1p_2 \tag{20.24a}$$

$$\tilde{\sigma}_y(\mathsf{x},\mathsf{p}) = x_1p_2 - p_1x_2 \tag{20.24b}$$

$$\tilde{\sigma}_z(\mathsf{x},\mathsf{p}) = \frac{1}{2}(x_1^2 + p_1^2 - x_2^2 - p_2^2)\ . \tag{20.24c}$$

We note that $\tilde{\sigma}_\alpha(\mathsf{x},\mathsf{p}) \equiv \sigma_\alpha^{\mathrm{cl}}(\mathsf{x},\mathsf{p})$ for $\alpha \in \{x,y,z\}$ and that the functional forms of σ_x^{w}, σ_y^{w} and σ_z^{w} are, therefore, simply given by their classical equivalent multiplied by an exponential factor. Of note is that the exponent is a conserved quantity of the mapping Hamiltonian.

The properties of the harmonic bath and its coupling to the quantum subsystem, namely the frequencies ω_j, masses M_j, and coefficients c_j are determined by the spectral density, $J(\omega)$, given by

$$J(\omega) = \frac{\pi}{2}\sum_j \frac{c_j^2}{M_j\omega_j}\delta(\omega-\omega_j)\ . \tag{20.25}$$

A common choice for the spectral density is for it to resemble the exponential form,

$$J(\omega) \approx \frac{\pi}{2}\ K\omega^s \exp(-\omega/\omega_c)\ , \tag{20.26}$$

where the Kondo parameter, K, and the characteristic frequency of the bath, ω_c, form the two key parameters determining the effect of the bath. When the exponent $s = 1$, the resulting spectral density is commonly referred to as having Ohmic character. This case corresponds to one of the most widely studied Spin-Boson problems, as the systems resulting from varying K and ω_c display a wide range of behaviors, from perfectly coherent to highly non-adiabatic [46, 49–51].

For any given Ohmic Spin-Boson problem, once the Kondo parameter and characteristic frequency of the bath have been chosen, what remains is to sample from this distribution a set of F representative frequencies, ω_j, and corresponding coupling coefficients, c_j. The computational challenge of the underlying problem is highly dependent on the strategy chosen, as, for example, an abundance of high frequency bath modes, or modes with vanishing coefficients, can result in a significantly larger number of bath degrees of freedom being required to converge a given observable.

A number of different discretization schemes have been proposed [52–55]. In this work, we select them according to the density

$$\rho(\omega_j) = F/\omega_c \exp(-\omega_j/\omega_c)\ , \tag{20.27}$$

such that $j = \int_0^{\omega_j}\rho(\omega)\mathrm{d}\omega$. This gives us the couplings

$$c_j = \sqrt{\frac{2M\omega_j J(\omega_j)}{\pi\rho(\omega_j)}} = \omega_j\sqrt{MK\omega_c/F}\ , \tag{20.28}$$

and the frequencies

$$\omega_j = -\omega_c \ln\left(\frac{\frac{1}{2}+j}{F}\right)\ . \tag{20.29}$$

This particular discretization scheme has the advantage that it reproduces the exact value for the reorganization energy, $2\sum_j c_j^2/M\omega_j^2 = 2K\omega_c$, for any number of bath modes.

The Spin-Boson Hamiltonian can be used to model a number of different physical processes, including electron transfer reactions, proton tunneling as well as a number of photo-induced phenomena [41–43]. We are, however, also interested in more general molecular systems which cannot be accurately described by sets of harmonic oscillators. For this reason, in the following, we will not assume that the Hamiltonian has the Spin-Boson form and we will discuss methods based on classical trajectories which can be applied equally well to more realistic anharmonic systems. The Spin-Boson model is however a convenient system for benchmarking their accuracy.

20.3.2 Non-Equilibrium Correlation Functions

As mentioned in Section 20.1, most properties of interest related to the study of complex systems are defined in terms of time correlation functions. Here we will consider non-equilibrium correlation functions which are initially prepared in an uncoupled thermal bath state. At time $t = 0$, the coupling is switched on and the dynamics are initiated.

The thermal density matrix of the bath is given by $\hat{\rho}_b = \mathrm{e}^{-\beta \hat{H}_b}/Z_b$ where $Z_b = \mathrm{Tr}_b[\mathrm{e}^{-\beta \hat{H}_b}]$ is a trace over the bath modes only. Because there are two electronic states, $\mathrm{Tr}[\hat{\rho}_b] = 2$, so the initial density matrix is defined as $\hat{\rho} = \frac{1}{2}\hat{\rho}_b$, such that $\mathrm{Tr}[\hat{\rho}] = 1$.

Sampling initial conditions for the nuclear (bath) degrees of freedom is straightforward for this system, as it is simply a set of uncoupled harmonic oscillators. The thermal Wigner distribution, defined by the Wigner transform discussed above, can be used to straightforwardly sample phase-space variables for a quantum system. In the case of the Spin-Boson Hamiltonian, this distribution for the bath is

$$\rho_b^{\mathrm{w}}(\vec{R}, \vec{P}) = \prod_{j=1}^{F} 2 \tanh\left(\frac{1}{2}\beta\hbar\omega_j\right) \exp\left[-\frac{1}{\hbar}\left(\frac{P_j^2}{M_j\omega_j} + M_j\omega_j R_j^2\right)\tanh\left(\frac{1}{2}\beta\hbar\omega_j\right)\right], \tag{20.30}$$

which is normalized such that $\frac{1}{(2\pi\hbar)^F}\int \rho_b^{\mathrm{w}}(\vec{R}, \vec{P})\mathrm{d}\vec{R}\mathrm{d}\vec{P} = 1$.

As discussed in section Section 20.2.2, the definition of the LSC-IVR correlation function involves Wigner transforms of the correlated operators. For this Hamiltonian, and in fact any system written in terms of the Pauli matrices, sampling the electronic degrees of freedom, i.e., the mapping variables x and p is, therefore, somewhat more involved, owing to the different ways to Wigner transform the Pauli operators, discussed above.

Given the different results of the Wigner transform, based on the representation chosen for the Pauli matrix operators, there are two strategies for turning the operators $\hat{A}$ and $\hat{B}$ into functions of the extended phase-space variables, allowing them to be calculated from semi-classical trajectories. Therefore, two different types of LSC-IVR correlation functions of Pauli matrices are possible.

The first uses two Wigner transformed operators, as given in Eqs. (20.23), which corresponds to two projections onto the harmonic oscillator subspace, one for each operator. Because the exponential term is conserved by the dynamics, one can sample the initial values of the mapping variables from a function, $\rho_s^{(2)}(\mathsf{x}, \mathsf{p})$, corresponding to the square of the exponential factor shared by A^{w} and B^{w}, namely

$$\rho_s^{(2)}(\mathsf{x}, \mathsf{p}) = (2/\pi)^2\ \mathrm{e}^{-2(x_1^2+p_1^2+x_2^2+p_2^2)}\ , \tag{20.31}$$

where the superscript (2) represents the two projections carried out for the operators. After sampling from this distribution, the values of the Pauli operators are calculated from the mapping

variables using the expressions in Eqs. (20.23). The correlation function is, therefore, given by

$$C_{AB}(t) \approx \frac{1}{(2\pi\hbar)^F(2\pi)^2} \iiiint \frac{1}{2}\rho_b^{\mathrm{w}}(\vec{R},\vec{P})A^{\mathrm{w}}(\mathsf{x}(0),\mathsf{p}(0))B^{\mathrm{w}}(\mathsf{x}(t),\mathsf{p}(t))\,\mathrm{d}\mathsf{x}\,\mathrm{d}\mathsf{p}\,\mathrm{d}\vec{R}\,\mathrm{d}\vec{P} \tag{20.32a}$$

$$= 8\langle \tilde{A}(\mathsf{x}(0),\mathsf{p}(0))\,\tilde{B}(\mathsf{x}(t),\mathsf{p}(t))\rangle_{\rho_b^{\mathrm{w}}\rho_s^{(2)}}\,, \tag{20.32b}$$

where $\hat{A}$ and $\hat{B}$ have been assumed to be pure subsystem operators. The angular brackets imply a classical ensemble average,

$$\langle\mathcal{O}\rangle_\rho = \frac{\iiiint \rho\mathcal{O}\,\mathrm{d}\mathsf{x}\,\mathrm{d}\mathsf{p}\,\mathrm{d}\vec{R}\,\mathrm{d}\vec{P}}{\iiiint \rho\,\mathrm{d}\mathsf{x}\,\mathrm{d}\mathsf{p}\,\mathrm{d}\vec{R}\,\mathrm{d}\vec{P}}\,, \tag{20.33}$$

and can be computed from an average of the value of $\mathcal{O}$ obtained from random sample of the distribution ρ.

Intuitively, one sets $\tilde{A} = 1$ in order to obtain an expression for the time-dependent observable averages,

$$P_B(t) \approx \frac{1}{(2\pi\hbar)^F(2\pi)^2} \iiiint \frac{1}{2}\rho_b^{\mathrm{w}}(\vec{R},\vec{P})B^{\mathrm{w}}(\mathsf{x}(t),\mathsf{p}(t))\,\mathrm{d}\mathsf{x}\,\mathrm{d}\mathsf{p}\,\mathrm{d}\vec{R}\,\mathrm{d}\vec{P} \tag{20.34a}$$

$$= 8\langle \tilde{B}(\mathsf{x}(t),\mathsf{p}(t))\rangle_{\rho_b^{\mathrm{w}}\rho_s^{(2)}}. \tag{20.34b}$$

A different approximation to the correlation function is obtained if only one of the operators is projected onto the harmonic oscillator subspace before Wigner transforming, the other being taken to be of the classical form shown in Eqs. (20.22). This second definition differs only in the number of exponential prefactors present in the phase-space representation of the operator. The same prescription for calculating $C_{AB}(t)$ can therefore be followed, if the mapping variables are sampled from

$$\rho_s^{(1)}(\mathsf{x},\mathsf{p}) = \pi^{-2}\mathrm{e}^{-(x_1^2+p_1^2+x_2^2+p_2^2)}\,, \tag{20.35}$$

where the superscript (1) represents a single projection onto the harmonic oscillator subspace. This second definition of the correlation function is therefore

$$C_{AB}(t) \approx \frac{1}{(2\pi\hbar)^F(2\pi)^2} \iiiint \frac{1}{2}\rho_b^{\mathrm{w}}(\vec{R},\vec{P})A^{\mathrm{cl}}(\mathsf{x}(0),\mathsf{p}(0))B^{\mathrm{w}}(\mathsf{x}(t),\mathsf{p}(t))\,\mathrm{d}\mathsf{x}\,\mathrm{d}\mathsf{p}\,\mathrm{d}\vec{R}\,\mathrm{d}\vec{P} \tag{20.36a}$$

$$= 2\langle \tilde{A}(\mathsf{x}(0),\mathsf{p}(0))\tilde{B}(\mathsf{x}(t),\mathsf{p}(t))\rangle_{\rho_b^{\mathrm{w}}\rho_s^{(1)}}\,. \tag{20.36b}$$

We would have obtained an equivalent expression using A^{w} and B^{cl}. Choosing $\tilde{A} = 1$,

$$P_B(t) \approx \frac{1}{(2\pi\hbar)^F(2\pi)^2} \iiiint \frac{1}{2}\rho_b^{\mathrm{w}}(\vec{R},\vec{P})B^{\mathrm{w}}(\mathsf{x}(t),\mathsf{p}(t))\mathrm{d}\mathsf{x}\,\mathrm{d}\mathsf{p}\,\mathrm{d}\vec{R}\,\mathrm{d}\vec{P} \tag{20.37a}$$

$$= 2\langle \tilde{B}(\mathsf{x}(t),\mathsf{p}(t))\rangle_{\rho_b^{\mathrm{w}}\rho_s^{(1)}}. \tag{20.37b}$$

A third approach, which does not project either operator onto the harmonic oscillator subspace, is clearly not allowed. This would not be physically relevant as it would not map back onto the real non-adiabatic problem. Furthermore, in the absence of projected operators, there exists no distribution to sample the mapping variables from, as the exponential distributions in Eq. (20.31) and Eq. (20.35) arise from the Wigner transforms of the projected operators.

To illustrate the difference between the two approximations for the correlation functions and expectation values introduced above we have calculated the Pauli operator correlation functions and expectation values for a biased Spin-Boson Hamiltonian, the results and parameters of which

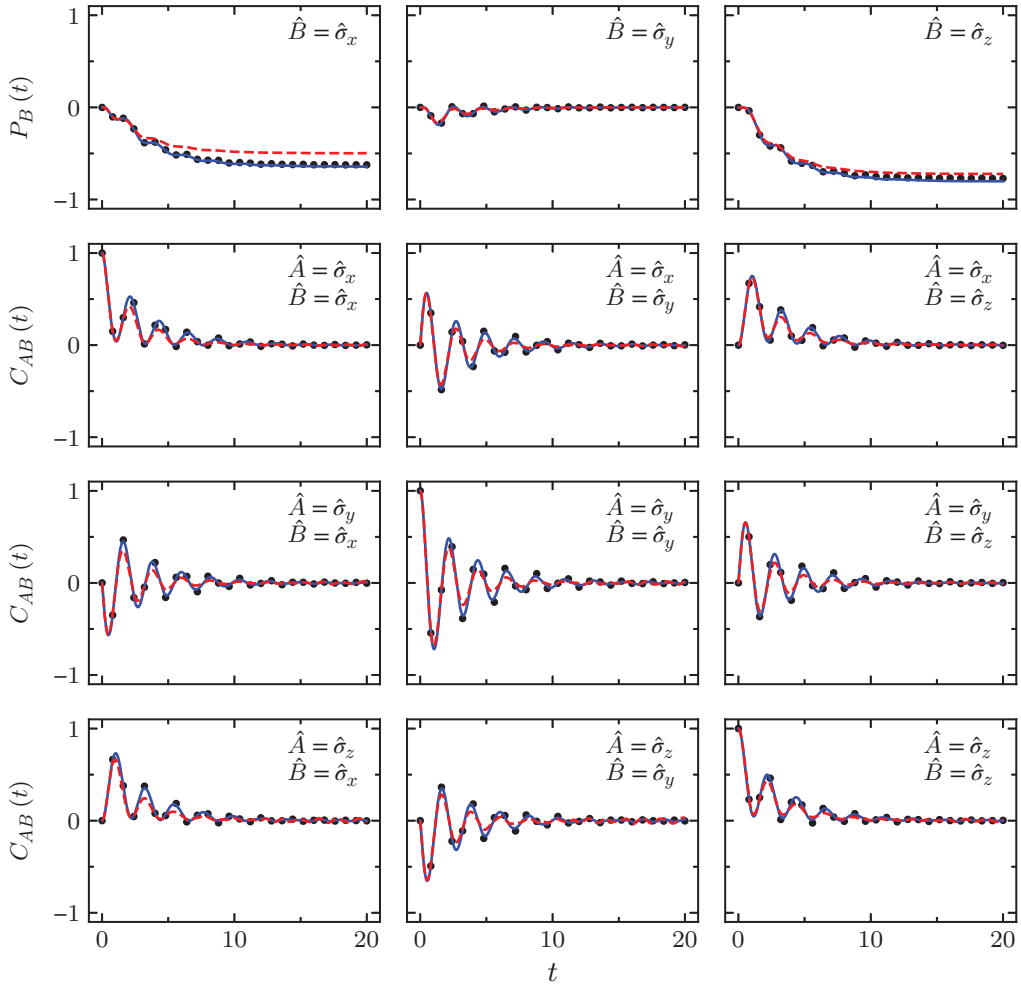


Figure 20.2 Pauli operator correlation functions and expectation values from LSC-IVR for a Spin-Boson model with energy bias $\epsilon = \Delta$, bath characteristic frequency $\omega_c = 2.5\Delta$ and Kondo parameter $K = 0.2$, at inverse temperature $\beta = 10\Delta^{-1}$ [56]. The solid blue line corresponds to a doubly projected correlation function, as per Eq. (20.32), the red dashed line to the singly projected version, shown in Eq. (20.36). Numerically exact QUAPI results [56] are shown as black dots.

are given in Figure 20.2. Owing to the efficient discretization scheme chosen for the bath degrees of freedom, only $F = 36$ bath modes were required to reach convergence with respect to numerically exact quasi-adiabatic propagator path integral (QUAPI) results. Operators were calculated as averages over a total of 10^7 semiclassical trajectories. Comparing the doubly (solid blue lines) and singly projected (red dashed lines) correlation functions and expectation values, it becomes clear that the difference is relatively small, all major structural features being present in both. A more detailed comparison of the two formulations can be found in Ref. [57].

The correlation functions display a number of highly symmetrical properties. In cases where $\hat{A} = \hat{B}$ for example, the initial value is $C_{AB}(0) = 1$, while in all other cases where $\hat{A} \neq \hat{B}$, it is $C_{AB}(0) = 0$. These values in the short time limit are exact and retained by the approximation, as they analytically derive from the underlying algebra of the Pauli matrices. Furthermore, at times $t > 0$, there

exists a set of symmetry relationships where

$$C_{\sigma_x\sigma_y}(t) = -C_{\sigma_y\sigma_x}(t) \tag{20.38a}$$

$$C_{\sigma_z\sigma_y}(t) = -C_{\sigma_y\sigma_z}(t) \tag{20.38b}$$

$$C_{\sigma_x\sigma_z}(t) = \quad C_{\sigma_z\sigma_x}(t)\ . \tag{20.38c}$$

These symmetries are also obeyed by both approximations.

When using the Spin-Boson Hamiltonian to model processes such as electron transfer, it is typical to study the dynamics of a system initialized in one of the states and to calculate the population difference between the electronic states as a function of time. The population difference is measured by the σ_z operator but the initial conditions are different from those employed above. Although we have not directly computed the function of interest, we are able to take linear combinations of our correlation functions to obtain the required result.

In the notation of the Pauli matrices, the initial condition corresponds to $|1\rangle\langle 1|$, which can, in terms of the Pauli matrices, be constructed as follows

$$|1\rangle\langle 1| \equiv \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \frac{1}{2}\left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\right] = \frac{1}{2}(\hat{I} + \hat{\sigma}_z)\ . \tag{20.39}$$

Therefore the function of interest can be calculated as follows:

$$\mathrm{Tr}[\hat{\rho}_b|1\rangle\langle 1|\hat{\sigma}_z(t)] \equiv C_{I\sigma_z}(t) + C_{\sigma_z\sigma_z}(t) \equiv P_{\sigma_z}(t) + C_{\sigma_z\sigma_z}(t)\ . \tag{20.40}$$

The second equivalency holds in the exact quantum picture, but not in the semiclassical approximation. Refs. 56 and 57 show computing the correlation functions in this way significantly improves the quality of semiclassical dynamics results for population differences. Note that the factor of half in Eq. (20.39) is included in Eq. (20.40) as part of the normalization of the correlation functions.

Figure 20.3 shows the population difference, with the system starting in the $|1\rangle$ state at $t = 0$, calculated using both doubly (solid blue line) and singly projected operators (red dashed line), as well as numerically exact results from a grid based QUAPI calculation [56]. Overall, this example serves to show how well LSC-IVR combined with the mapping approach can describe the dynamics of the Spin-Boson model. Again the difference between the two strategies for computing correlation functions is rather small, the double projection definition being in slightly better agreement with the exact result.

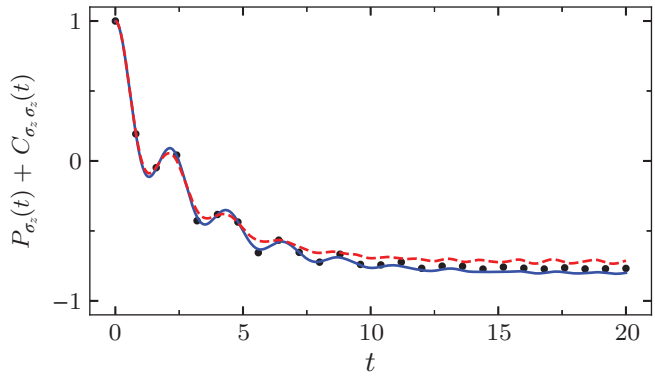


Figure 20.3 Population difference starting from electronic state $|1\rangle$, computed from LSC-IVR correlation functions and expectation values. Solid blue and red dashed lines correspond to doubly and singly projected correlation functions respectively, black dots to numerically exact QUAPI results [56].

20.4 Non-Adiabatic Path-Integral Theory

In the previous sections, we have discussed the use of linearized path-integral methods for simulating the dynamics of non-adiabatic systems. Such approaches employ purely classical dynamics for the nuclear degrees of freedom and are combined in various ways with a quantum description of the electronic-state dynamics. In this section, we introduce imaginary-time path-integral methods which allow us to include quantum effects in both the electronic and nuclear degrees of freedom.

These methods are particularly useful for simulating dynamics initialized from an equilibrium distribution, as the imaginary-time path-integral methods sample this naturally. Much progress has been made with these techniques for studying single-PES adiabatic systems in order to obtain thermodynamic information on condensed-phase systems [59–62]. Real-time correlation functions can also be obtained using ring-polymer molecular dynamics [3, 63]. This is an approximate method, but as it is also based on the imaginary-time path-integral, it can be shown that it tends to the exact result in the short-time limit. A number of non-adiabatic extensions of RPMD have been suggested based on the mapping approach [64–66]. We will introduce the idea below and show benchmark results from this methodology.

20.4.1 Mean-Field Path-Integral Sampling

First we will give a derivation of the imaginary-time path-integral approach for computing thermodynamic quantities in non-adiabatic systems [67]. This, like the original path-integral methodology [1], can be derived from a splitting of the quantum Boltzmann operator.

The total partition function is the trace of the Boltzmann operator,

$$Z = \mathrm{Tr}[\mathrm{e}^{-\beta\hat{H}}], \tag{20.41}$$

and can be split into factors depending alternately on the potential or kinetic energy operators

$$Z \simeq \mathrm{Tr}\left[\prod_{i=1}^{N} \mathrm{e}^{-\beta_N(\hat{U}+\hat{V})/2}\mathrm{e}^{-\beta_N\hat{\vec{P}}^2/2M}\mathrm{e}^{-\beta_N(\hat{U}+\hat{V})/2}\right], \tag{20.42}$$

where $\beta_N = \beta/N$. This expression is exact in the $N \to \infty$ limit where β_N is small.

We now insert a complete set of position states, $1 = \int |\vec{R}_i\rangle \langle\vec{R}_i| \mathrm{d}\vec{R}_i$, for each value of i to evaluate the potential operators, $\mathrm{e}^{-\beta_N(\hat{U}+\hat{V})/2}|\vec{R}_i\rangle = \mathrm{e}^{-\beta_N(U(\vec{R}_i)+V(\vec{R}_i))/2}|\vec{R}_i\rangle$. Similarly we evaluate the kinetic operators by inserting complete sets of momentum states $1 = \int |\vec{P}_i\rangle \langle\vec{P}_i| \mathrm{d}\vec{P}_i$. The momentum variables can be integrated out analytically because the integrand is Gaussian. In the case of a single potential energy surface, $V \equiv 0$, this yields the expression

$$Z \simeq \left(\frac{M}{2\pi\beta_N\hbar^2}\right)^{NF/2} \int \mathrm{e}^{-\beta_N U_N(\boldsymbol{R})} \mathrm{d}\boldsymbol{R}, \tag{20.43}$$

where $\boldsymbol{R} = \{\vec{R}_1, \ldots, \vec{R}_N\}$, F is the number of nuclear degrees of freedom, and

$$U_N(\boldsymbol{R}) = \sum_{i=1}^{N}\left(\frac{M}{2\beta_N^2\hbar^2}|\vec{R}_i - \vec{R}_{i-1}|^2 + U(\vec{R}_i)\right), \tag{20.44}$$

with cyclic indices such that $\vec{R}_0 \equiv \vec{R}_N$, is an effective potential. We will call it the *ring-polymer potential* because one can interpret it as the potential of a classical ring-polymer with N beads connected

by harmonic springs. This can be generalized to a multi-surface system with diabatic potential matrix V in which case we instead have

$$Z \simeq \left(\frac{M}{2\pi\beta_N\hbar^2}\right)^{NF/2} \int W_0(\boldsymbol{R}) \mathrm{e}^{-\beta_N U_N(\boldsymbol{R})} \mathrm{d}\boldsymbol{R}, \tag{20.45}$$

with

$$W_0(\boldsymbol{R}) = \mathrm{tr}\left[\prod_{i=1}^{N} \mathrm{e}^{-\beta_N \mathsf{V}(\vec{R}_i)}\right], \tag{20.46}$$

where the notation tr[•] refers to a trace over only the electronic degrees of freedom.

In order to compute quantum equilibrium statistical averages of an operator $\hat{A} = A(\hat{\vec{R}})$, a similar derivation gives

$$\langle \hat{A} \rangle_{\mathrm{eq}} = \frac{1}{Z} \mathrm{Tr}[A(\hat{\vec{R}}) \mathrm{e}^{-\beta \hat{H}}] \tag{20.47a}$$

$$\simeq \frac{\int \bar{A}(\boldsymbol{R}) \rho_0(\boldsymbol{R}) \mathrm{d}\boldsymbol{R}}{\int \rho_0(\boldsymbol{R}) \mathrm{d}\boldsymbol{R}} \equiv \langle \bar{A}(\boldsymbol{R}) \rangle_{\rho_0}, \tag{20.47b}$$

where $\rho_0(\boldsymbol{R}) = W_0(\boldsymbol{R}) \mathrm{e}^{-\beta_N U_N(\boldsymbol{R})}$ and $\bar{A}(\boldsymbol{R}) = \frac{1}{N}\sum_i A(\vec{R}_i)$. Here we have used the fact that all beads are equivalent to define the statistical average in a symmetric way. This will reduce the statistical error from a given sample.

In order to compute the average over a state-dependent operator, $\hat{A} = |\nu\rangle \langle\mu|$, a projection onto the required states is inserted into the trace over matrix exponentials. This gives

$$\langle \hat{A} \rangle_{\mathrm{eq}} = \frac{1}{Z} \mathrm{Tr}[|\nu\rangle \langle\mu| \mathrm{e}^{-\beta \hat{H}}] \tag{20.48a}$$

$$\simeq \frac{\int [\Gamma_{\nu\mu}(\boldsymbol{R})/W_0(\boldsymbol{R})] \rho_0(\boldsymbol{R}) \mathrm{d}\boldsymbol{R}}{\int \rho_0(\boldsymbol{R}) \mathrm{d}\boldsymbol{R}} \equiv \left\langle \frac{\Gamma_{\nu\mu}(\boldsymbol{R})}{W_0(\boldsymbol{R})} \right\rangle_{\rho_0}, \tag{20.48b}$$

where here $\Gamma_{\nu\mu}(\boldsymbol{R}) = \left[\prod_{i=1}^{N} \mathrm{e}^{-\beta_N \mathsf{V}(\vec{R}_i)}\right]_{\nu\mu}$. A similar symmetrization procedure could also be applied in this case, since the trace is cyclically permutable.

These statistical averages can be evaluated either by path-integral Monte Carlo techniques [68] or by path-integral molecular dynamics using an effective potential of $U_N(\boldsymbol{R}) - \frac{1}{\beta_N} \ln W_0(\boldsymbol{R})$. We emphasize that such dynamics is just a sampling tool and *cannot* be used to calculate real-time dependence, in contrast to what will be presented in section 20.4.2. It may be noted that sometimes the estimator $\bar{A}$ at zero time is better incorporated by modifying the distribution than by an average over beads and mapping variables [69]. Extensions to deal with operators of momenta can be obtained in the usual way [4].

20.4.2 Non-Adiabatic Ring-Polymer Molecular Dynamics

In Section 20.4.1, we introduced a non-adiabatic imaginary-time path-integral formulation which can be used to compute statistical properties of systems with multiple electronic states. However, such a formulation cannot be directly employed to obtain information on real-time dynamics as it has no memory for the populations of the diabatic states and treats them only in an average way. In Section 20.2.2 we already explored one method to approximate real-time dynamics, LSC-IVR, which has the disadvantage that it requires us to compute a Wigner transform, which is only possible for simple potentials. To allow initial distributions on general potential energy surfaces, we will make use of the path-integral formalism.

First we briefly present how to generate real-time dynamics on a single potential energy surface ($\mathsf{V} \equiv 0$) using ring-polymer molecular dynamics (RPMD). It has been shown that if one generates equations of motion from the Hamiltonian

$$\mathcal{H}_N^{\mathrm{RPMD}} = \frac{\boldsymbol{P}^2}{2M} + U_N(\boldsymbol{R}), \tag{20.49}$$

the dynamics can be used to approximate Kubo-transformed correlation functions $\tilde{C}_{AB}(t)$, defined in Eq. (20.6). The correlation functions are exact at $t = 0$, accurate to an order of at least $\mathcal{O}(t^4)$ for general $A(\hat{\vec{R}})$, $B(\hat{\vec{R}})$, and reduces to classical dynamics at high temperature [63]. Compared to a classical treatment, RPMD includes nuclear quantum effects by treating the nuclei in a delocalized way.

In order to get real-time information from ring-polymer dynamics on multiple surfaces (general V), we follow the mapping approach and represent the discrete electronic levels by continuous phase-space variables. In particular we will present the approach called non-adiabatic ring-polymer molecular dynamics (NRPMD) [64].

The NRPMD Hamiltonian is defined as [9, 64]

$$\mathcal{H}_N^{\mathrm{NRPMD}} = \frac{\boldsymbol{P}^2}{2M} + U_N(\boldsymbol{R}) + \sum_{i=1}^{N} \frac{1}{2}[\mathsf{x}_i^{\mathrm{T}}\mathsf{V}_i\mathsf{x}_i + \mathsf{p}_i^{\mathrm{T}}\mathsf{V}_i\mathsf{p}_i - \mathrm{tr}\ \mathsf{V}_i], \tag{20.50}$$

where $\boldsymbol{x} = \{\mathsf{x}_1, \dots, \mathsf{x}_N\}$ and $\boldsymbol{p} = \{\mathsf{p}_1, \dots, \mathsf{p}_N\}$ are sets of mapping variables assigned to each ring-polymer bead, and $\mathsf{V}_i = \mathsf{V}(\vec{R}_i)$. The first two terms are the same as in standard (single surface) RPMD, and the third term adds contributions from different potential energy surfaces. From this Hamiltonian we obtain the equations of motion

$$\dot{\vec{R}}_i = \frac{\vec{P}_i}{M} \tag{20.51a}$$

$$\dot{\vec{P}}_i = -\frac{\partial U}{\partial \vec{R}_i} - \frac{1}{2}\left[\mathsf{x}_i^{\mathrm{T}}\frac{\partial \mathsf{V}_i}{\partial \vec{R}_i}\mathsf{x}_i + \mathsf{p}_i^{\mathrm{T}}\frac{\partial \mathsf{V}_i}{\partial \vec{R}_i}\mathsf{p}_i - \mathrm{tr}\ \frac{\partial \mathsf{V}_i}{\partial \vec{R}_i}\right] \tag{20.51b}$$

$$\dot{\mathsf{x}}_i = \mathsf{V}_i\mathsf{p}_i \tag{20.51c}$$

$$\dot{\mathsf{p}}_i = -\mathsf{V}_i\mathsf{x}_i. \tag{20.51d}$$

In the ring-polymer molecular dynamics ansatz, one uses the dynamics of such a Hamiltonian to compute real-time correlation functions. The dynamics of these trajectories will only be approximate, but can still give reasonable results on short time scales, which are the most important in many applications. The initial conditions for the trajectories should be sampled from the quantum Boltzmann distribution, written in a path-integral form.

We cannot directly use the distribution obtained in Section 20.4.1 as this would not give a distribution for the mapping variables. Therefore, instead of computing the partition function in a mean-field way, we will express it in terms of mapping variables. This gives us an equilibrium distribution similar to the one derived by Ananth and Miller [70]. Extending the derivation in Section 20.4.1 by also inserting complete sets of eigenstates of the operators $\hat{\mathsf{x}}_i$ and $\hat{\mathsf{p}}_i$, and further factorizing each factor $\mathrm{e}^{-\beta_N \mathsf{V}(\vec{R}_i)}$ into two matrices, we can write [9, 64]

$$W_1(\boldsymbol{R}, \boldsymbol{x}, \boldsymbol{p}) = \mathrm{tr}\left[\prod_{i=1}^{N}\sum_{\alpha,\beta,\gamma,\delta=1}^{L} \mathrm{e}^{-\beta_N \mathsf{V}(\vec{R}_i)/2}|\alpha\rangle\langle\alpha|\mathsf{x}_i\rangle\langle\mathsf{x}_i|\beta\rangle\langle\beta|\mathrm{e}^{-\beta_N \mathsf{V}(\vec{R}_i)/2}|\gamma\rangle\langle\gamma|\mathsf{p}_i\rangle\langle\mathsf{p}_i|\delta\rangle\langle\delta|\right]. \tag{20.52}$$

The factors $\sum_{\alpha=1}^{L}|\alpha\rangle\langle\alpha| \equiv \sum_{\alpha=1}^{L}|0_1, \dots, 1_\alpha \dots, 0_L\rangle\langle 0_1, \dots, 1_\alpha, \dots, 0_L|$ are required to project the mapping variables on to the physical subspace of singly-excited oscillators [70].

The total partition function can now be written in the form

$$Z \simeq \frac{1}{(2\pi\hbar)^{NF}} \int \rho_1(\boldsymbol{R},\boldsymbol{P},\boldsymbol{x},\boldsymbol{p})\, \mathrm{d}\boldsymbol{R}\, \mathrm{d}\boldsymbol{P}\, \mathrm{d}\boldsymbol{x}\, \mathrm{d}\boldsymbol{p}, \tag{20.53}$$

where the distribution is

$$\rho_1(\boldsymbol{R},\boldsymbol{P},\boldsymbol{x},\boldsymbol{p}) = \left(\frac{4}{\pi^2}\right)^N W_1(\boldsymbol{R},\boldsymbol{x},\boldsymbol{p}) \mathrm{e}^{-\boldsymbol{x}^2-\boldsymbol{p}^2} \mathrm{e}^{-\beta_N[\boldsymbol{P}^2/2m+U_N(\boldsymbol{R})]}\,, \tag{20.54}$$

and

$$W_1(\boldsymbol{R},\boldsymbol{x},\boldsymbol{p}) = \mathrm{tr}[\Gamma(\boldsymbol{R},\boldsymbol{x},\boldsymbol{p})]\,. \tag{20.55}$$

The state-independent potentials $U(\vec{R}_i)$ are included in $U_N(\boldsymbol{R})$, Eq. (20.44), whereas the potential matrices, $\mathsf{V}(\vec{R}_i)$, are in

$$\Gamma(\boldsymbol{R},\boldsymbol{x},\boldsymbol{p}) = \prod_{i=1}^{N} \mathsf{M}_i \mathsf{x}_i \mathsf{x}_i^{\mathrm{T}} \mathsf{M}_i \mathsf{p}_i \mathsf{p}_i^{\mathrm{T}}, \tag{20.56}$$

through the matrices

$$\mathsf{M}_i = \mathrm{e}^{-\beta_N \mathsf{V}(\vec{R}_i)/2}. \tag{20.57}$$

Thermodynamic statistical quantities can be obtained similar to the mean-field approach described in Section 20.4.1, and, in the limit of $N \to \infty$, these will tend to the exact quantum result. The distribution ρ_1 can be sampled in a Monte Carlo scheme using trial moves for both nuclear and electronic coordinates. It is important to notice that W_1 (and thereby also ρ_1) in general is not positive definite. When sampling the thermal average of an observable $\mathcal{O}$ one therefore needs to weight each sample by its sign

$$\langle \mathcal{O} \rangle_{\rho_1} = \frac{\langle \mathcal{O} \mathrm{sgn}\ \rho_1 \rangle_{|\rho_1|}}{\langle \mathrm{sgn}\ \rho_1 \rangle_{|\rho_1|}}. \tag{20.58}$$

For high N this can decrease the sampling efficiency and we will discuss a way to alleviate this problem in Section 20.4.3. Therefore, for thermodynamic quantities, one should use the mean-field approach and the advantage of the distribution ρ_1 is that it can be used to sample initial values for real time trajectories.

As in the standard RPMD ansatz [3], this method can be used to approximate Kubo-transformed correlation functions, Eq. (20.6). In NRPMD this is computed from trajectories initialized according to the distribution ρ_1,

$$\tilde{C}_{AB}(t) \approx \langle \bar{A}(\boldsymbol{R}(0),\boldsymbol{x}(0),\boldsymbol{p}(0)) \bar{B}(\boldsymbol{R}(t),\boldsymbol{x}(t),\boldsymbol{p}(t)) \rangle_{\rho_1}, \tag{20.59}$$

where $\bar{B} = \frac{1}{N}\sum_i^N B(\vec{R}_i, \mathsf{x}_i, \mathsf{p}_i)$ and B is obtained by replacing the quantum operators in the mapping representation of $\hat{B}$ by classical coordinates [9]. The operator $\bar{A}$ is defined in a similar way if it is an operator of nuclear positions. However, for an operator of the form $\hat{A} = |\nu\rangle\langle\mu|$, the explicit expression is

$$\bar{A} = \frac{\Gamma_{\nu\mu}(\boldsymbol{R},\boldsymbol{x},\boldsymbol{p})}{W_1(\boldsymbol{R},\boldsymbol{x},\boldsymbol{p})}\,, \tag{20.60}$$

or alternatively

$$\bar{A} = \frac{1}{N}\sum_{i=1}^{N} \frac{[\mathsf{p}_{i-1}]_\nu [\mathsf{M}_i \mathsf{x}_i]_\mu}{\mathsf{p}_{i-1}^{\mathrm{T}} \mathsf{M}_i \mathsf{x}_i}, \tag{20.61}$$

where we have used the cyclicity of the trace to symmetrize over beads and mapping variables.

It should be mentioned that the method presented above is not the only way to unite mapping variables and RPMD. A similar approach by Chowdhury and Huo [66], Coherent State-RPMD (CS-RPMD), is based on a coherent state mapping basis and uses the same Hamiltonian for the dynamics, but has a different initial distribution than in the NRPMD approach. The so-called mapping-variable RPMD (MV-RPMD) by Ananth [65] uses a different Hamiltonian for the dynamics designed to conserve the Boltzmann distribution, which is not generally true for NRPMD and CS-RPMD. Unfortunately this comes with the disadvantage that it cannot exactly describe Rabi oscillations in an uncoupled two-level system, which is possible with NRPMD and CS-RPMD.

20.4.3 Alleviation of the Negative Sign

It is possible to carry out a canonical transformation to the action-angle coordinates J_{in}, θ_{in},

$$x_{i\nu} = \sqrt{2J_{i\nu}+1}\sin\theta_{i\nu} \tag{20.62a}$$

$$p_{i\nu} = \sqrt{2J_{i\nu}+1}\cos\theta_{i\nu}. \tag{20.62b}$$

The action coordinates $J_{i\nu}$ thus have the form $\frac{1}{2}(x_{i\nu}^2 + p_{i\nu}^2 - 1)$ which is equivalent to the classical mapping function, Eq. (20.12a), of the population of state $|\nu\rangle$. We will therefore refer to $J_{i\nu}$ as population variables. In these coordinates the Hamiltonian has the form

$$\mathcal{H} = \frac{\boldsymbol{P}^2}{2M} + U_N(\boldsymbol{R}) + \sum_{i=1}^{N}\left[\sum_{\nu} J_{i\nu}V_\nu(\vec{R}_i) + \sum_{\nu\neq\mu} V_{\nu\mu}(\vec{R}_i)\sqrt{2J_{i\nu}+1}\sqrt{2J_{i\mu}+1}\cos(\theta_{i\nu}-\theta_{i\mu})\right]. \tag{20.63}$$

The Hamiltonian depends only on angular differences so that there is always an average angle $\Theta_i = \frac{1}{L}\sum_\nu \theta_{i\nu}$ of which the Hamiltonian is independent. This implies that the total population $\sum_\nu J_{i\nu}$ is conserved. It has been shown that it is possible to analytically integrate out this angular degree of freedom and that this alleviates the negative sign problem [9]. If one electronic level is uncoupled from the others, an additional angle can be integrated out. In the case of an uncoupled two-level system this means that both angles can be integrated out, which completely removes the sign problem for this system [9].

Nonetheless, in general, the distribution remains non-positive definitive. Special techniques developed to sample other non-positive definite distributions may also be applied in this case. One example is that of a set of fermions, for which a recent method was developed to sample both the positive and negative parts efficiently based on enhanced sampling with Metadynamics [71]. We note, however, that the fermionic sign problem is not necessarily equivalent to our sign problem, which appears to be much less problematic.

20.4.4 Practical Implementation of Monte Carlo Sampling

This section describes how to implement the Monte Carlo sampling of $\rho_1(\boldsymbol{R}, \boldsymbol{x}, \boldsymbol{p})$ in practice. We will use the generalized Metropolis algorithm [72], in which we generate trial moves according to an *a priori* sampling distribution $T(s \to s')$, that are accepted with probability $A(s \to s')$. A sufficient condition for equilibrium is to assert that the total transition probability $P(s \to s') = T(s \to s')$ $A(s \to s')$ satisfies *detailed balance*, which means that the average number of accepted moves from

s to s' is canceled out by the average number of reverse moves for any pair of two states s and s', i.e., $\rho(s)P(s \to s') = \rho(s')P(s' \to s)$, where $\rho(s)$ is the equilibrium distribution we want to sample. One can show that detailed balance holds if the acceptance probability is computed as

$$A(s \to s') = \min\left[1, \frac{T(s' \to s)\rho_1(s')}{T(s \to s')\rho_1(s)}\right]. \tag{20.64}$$

Here, s and s' are labels for the total state described by $\boldsymbol{R}$, $\boldsymbol{x}$ and $\boldsymbol{p}$.

There are many ways to generate trial moves. A simple and common type of move is to displace a coordinate with a random displacement vector sampled uniformly from within a cube of a given size. This is referred to as the classic move and can be used to update nuclear coordinates, $\vec{R}_i$, as well as mapping variables x_{iv}, p_{iv}. Typically, this move is too inefficient to sample the distribution on its own, and it is better to use a menu of different types of moves, where each move in the menu is used with a predefined probability. Rather than displacing just a single bead, we could as well displace the whole polymer (all beads) by the same vector, which is called a displacement move. Alternatively one can resample a nuclear bead coordinate between two neighboring beads according to a Gaussian distribution in the distance from their midpoint, a procedure called the free-particle move. This move will, in general, have a shorter correlation time than the classic move.

If the ring polymer has a large number of beads, it will be inefficient to use single-bead moves like the classic and free-particle moves. A faster way to generate new configurations is through normal-mode sampling, in which all the non-zero normal modes of the ring polymer are resampled from Gaussian distributions. For practical considerations of this and other nuclear sampling moves, the reader is referred to Ref. [68].

For the mapping variables we can add another set of trial moves. The simplest one, except for the classic move, is to generate the trial configuration from a Gaussian distribution

$$x_{iv} \sim (2\pi\sigma_x^2)^{-\frac{1}{2}} \exp\left(-\frac{x_{iv}^2}{2\sigma_x^2}\right) \tag{20.65a}$$

$$p_{iv} \sim (2\pi\sigma_p^2)^{-\frac{1}{2}} \exp\left(-\frac{p_{iv}^2}{2\sigma_p^2}\right), \tag{20.65b}$$

where σ_x and σ_p are typically chosen to be $1/\sqrt{2}$ but can be tuned to increase the average traveled distance per attempt. Another trial move is to swap the components of the mapping variables for a single bead: $(x_{iv}, x_{i\mu}) \to (x_{i\mu}, x_{iv})$ (and simultaneously the same operation for the p variables). This will increase the probability to cross between regions in phase space that are dominated by different potential energy surfaces. In a similar fashion, we can also let the mapping variables of different beads swap, $(x_{iv}, x_{i'v}) \to (x_{i'v}, x_{iv})$. Finally, the action-angle coordinates in Eqs. (20.62) allows for types of moves in which θ_{iv} are resampled uniformly from the interval $[0, 2\pi)$.

The correlation functions are evaluated using the same methods used in classical mechanics and standard RPMD. First, values of $\boldsymbol{R}_0$, $\boldsymbol{x}_0$ and $\boldsymbol{p}_0$ are selected from a Metropolis simulation, whereas $\boldsymbol{P}_0$ is chosen from a multidimensional normal distribution,

$$P_i \sim (2\pi M/\beta_N)^{-\frac{1}{2}} \exp(-\beta_N P_i^2/2M). \tag{20.66}$$

Trajectories are then run from these initial values using a velocity Verlet scheme, taking advantage of analytical updates for both the free ring-polymer normal modes and the harmonic mapping coordinate motions.

20.5 Equilibrium Correlation Functions

For illustration of the performance of the NRPMD method, we consider a system with two electronic states and a single vibrational mode, modeling, e.g., a vibrationally coupled electron-transfer process between a donor and acceptor state. The diabatic potentials are

$$U(R) = \frac{1}{2}M\omega^2 R^2 - \lambda(R) \tag{20.67a}$$

$$V_1(R) = +\alpha + \kappa R + \lambda(R) \tag{20.67b}$$

$$V_2(R) = -\alpha - \kappa R + \lambda(R)\,, \tag{20.67c}$$

where κ is the vibronic coupling and 2α the energy bias; the electronic coupling, Δ, is chosen constant. Reduced units are used such that M and $\hbar$ are equal to 1, and by setting the frequency $\omega = 1$, we effectively measure energy in units of ω. The function $\lambda(R) = \sqrt{(\alpha + \kappa R)^2 + \Delta^2}$ is chosen in order to force the lowest eigenvalue of $\mathrm{V}(R)$ to be 0 everywhere. This was found to improve the convergence of calculations.

In the separable case where $\kappa = 0$, the electronic and nuclear degrees of freedom are independent and the dynamics of the non-adiabatic RPMD method converges, in the large N limit, to the exact quantum result for the Kubo-transformed electronic correlation function,

$$\tilde{C}_{22}(t) = \frac{1+\gamma^2}{4} + \frac{\tanh\beta\hbar\Omega}{2}\left(\gamma + \frac{1-\gamma^2}{2\beta\hbar\Omega}\cos 2\Omega t\right), \tag{20.68}$$

where $\hbar\Omega = \sqrt{\alpha^2 + \Delta^2}$, $\gamma = \alpha/\hbar\Omega$ and the Rabi oscillation frequency is seen to be 2Ω.

The results of one such calculation are given in Figure 20.4. The correct Rabi frequency is clearly observed in both the classical ($N = 1$) and ring-polymer ($N = 4$) versions. However, unlike the computation of population probabilities [30], it is necessary, just as with semiclassical methods [70], to use the ring-polymer form in order to compute the correct amplitude of the thermal correlation function. These short-time results are presented in Table 20.1 and seen to tend quickly to the exact result.

To provide more stringent tests to the proposed method, correlation functions were computed for three non-separable systems and compared with numerically exact quantum results from a discrete-variable representation (DVR) calculation [73].

For strong electronic coupling, $\Delta = 4$, the timescale of the electronic oscillations is much shorter than that of the nuclear dynamics, such that the nuclei move in a mean field of the diabatic surfaces. This is close to the adiabatic limit where the nuclear dynamics tend to standard RPMD, which as shown in Figure 20.5(a), are almost exact in this case as $U(R)$ is approximately harmonic [3]. The proposed non-adiabatic RPMD method, converged with four beads, also provides an excellent approximation to the electronic correlation function, which only degrades slightly after a few oscillations, and is a significant improvement over the single-bead result. The reason why the method performs so well here is because of the choice $\lambda(R)$ such that the lowest eigenvalue of $\mathrm{V}(R)$ is zero, whereas the other is very high and unpopulated.

Table 20.1 Initial values of the electronic correlation function for the separable system with $\kappa = 0$, $\alpha = 0.5$ and $\Delta = 1$.

N	1	2	4	Exact
$\tilde{C}_{22}(0)$	0.639	0.629	0.625	0.625

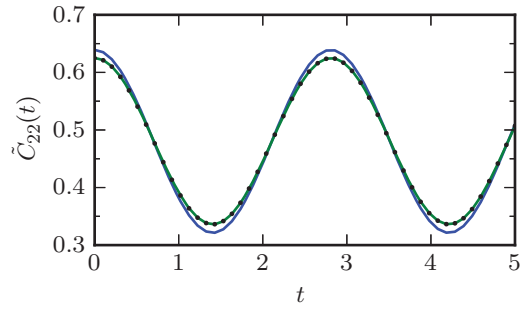


Figure 20.4 Calculated electronic correlation function, Eq. (20.68), for the separable system with $\kappa = 0$, $\alpha = 0.5$ and $\Delta = 1$ compared with the exact solution (black dots). The blue line shows the results with $N = 1$ and the green with $N = 4$.

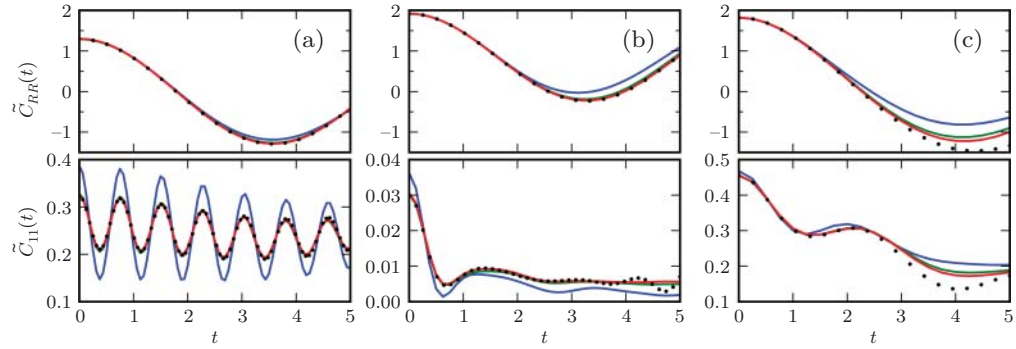


Figure 20.5 Calculated correlation functions, Eq. (20.59), for three systems with $\kappa = 1$ and $\beta = 1$: (a) close to the adiabatic limit with $\alpha = 0$ and $\Delta = 4$; (b) with a strong bias, $\alpha = 2$ and $\Delta = 1$; (c) in an intermediate regime with $\alpha = 0$ and $\Delta = 1$; compared with the DVR results (black dots). The blue, green, and red lines show the results of one, four, and eight bead calculations. Reproduced from Ref. [64].

With the parameters $\alpha = 2$ and $\Delta = 1$, the system enters the inverted Marcus regime. The major dynamical effects can be captured with the proposed method as shown in Figure 20.5(b), where again at least four beads are needed to describe them accurately. It is seen that the equilibrium population of the $|1\rangle$ state is much reduced in this case and is quickly lost to the lower-energy $|2\rangle$ state. The position autocorrelation function also shows that the system remains almost entirely on only one diabatic surface, but only with $N \geq 4$ is the method able to predict the correct amount of anharmonicity.

We finally consider a more challenging intermediate system with $\Delta = 1$ but without bias such that the timescales of the nuclear and electronic vibrations are similar. As seen in Figure 20.5(c), short-time results correct to within graphical accuracy are obtained with at least four beads. The approximation does, however, degrade after the first electronic oscillation, although this can be improved slightly using eight beads.

Finally, we note that, although exact quantum mechanics would predict a time-independent value of an observable when averaged over equilibrium conditions, this is not necessarily true of the non-adiabatic ring-polymer approach presented here. In order to test the behavior of the NRPMD approach, we present results in Figure 20.6 which show the deviation from the correct behavior. It appears that as N increases, the deviation becomes less troublesome. This may be one of the reasons why the NRPMD method is seen to give an improvement over the $N = 1$ version.

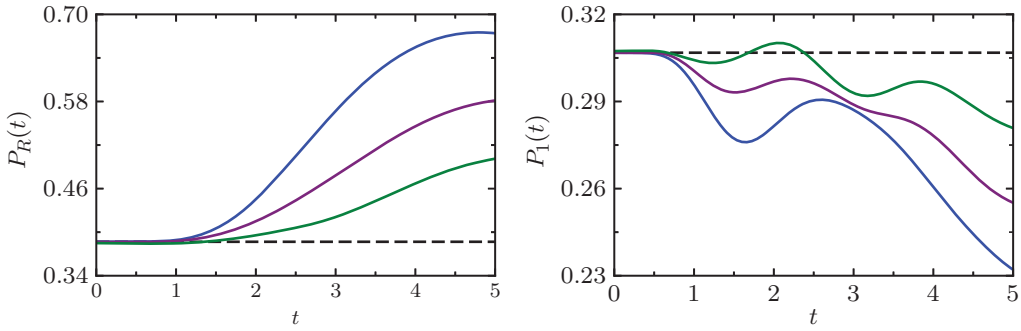


Figure 20.6 Calculated expectation values, Eq. (20.8), of the nuclear operator, $\hat{x}$, and the population of the electronic state $|1\rangle$, in the intermediate regime with $\alpha = 0.5$ and $\Delta = 1$. Blue, magenta, and green lines correspond to the results of calculations employing one, two, and four beads and the dashed line to the exact result.

20.6 Conclusions

We have presented two approaches based on path-integral methodology and the mapping representation for simulating non-adiabatic dynamics. These methods have been tested on simple model systems, but are also applicable directly to more complex and anharmonic descriptions of molecules.

LSC-IVR simulations compared very well with exact results, in particular the first version in which both operators use the projected Wigner transform. Results also show that the non-adiabatic RPMD method performs well in all regimes tested and tends to the exact, or at least standard RPMD, results in a number of limiting cases. It was also found to consistently improve upon the classical (single-bead) implementation of the mapping approach, presumably because of its accuracy in the short-time limit and better conservation of the harmonic oscillator subspace and Boltzmann distribution.

Nonetheless, both approaches are based on approximations which will break down under certain conditions. These methods will not provide correct long-time results in systems where nuclear quantum coherences play an important role, and are in fact only expected to be accurate at short times. We should expect NRPMD to be subject to the usual RPMD limitations such as causing spurious frequencies to appear in vibrational spectra [74, 75] and to have difficulties in describing correlation functions of strongly nonlinear operators [63].

Within a quasi-classical implementation of the mapping approach, problems with the flow of zero-point energy between the nuclear and electronic coordinates have been identified [52, 76]. In some circumstances, the population of a state can become negative, effectively inverting the potential [77, 78]. It is, however, expected that the NRPMD approach will alleviate these problems as the number of beads is increased [9].

The dynamics methods presented here can be used to obtain information about quantum statistics and dynamics of molecular systems and hence to determine observables such as rate constants and absorption spectra. Note, however, that in order to obtain rate constants, there may be more direct simulation routes based on approaches such as semiclassical instanton theory [79, 82], quantum instanton [12, 83, 84], and other forms of ring-polymer molecular dynamics [85, 86]. It may also be possible to combine these approaches with the mapping dynamics to develop further improved methods.

Acknowledgments

The authors acknowledge the financial support from the Swiss National Science Foundation through the NCCR MUST (Molecular Ultrafast Science and Technology) Network. M.A.C.S. would like to acknowledge the support of this work through the award of the ETH postdoctoral fellowship by the Swiss Federal Institute of Technology in Zurich. J.E.R. would like to acknowledge support through the Hans H. Günthard Doctoral Fellowship of the same institute.

References

1 Feynman, R.P. and Hibbs, A.R. (1965). *Quantum Mechanics and Path Integrals*. New York: McGraw-Hill.
2 Kubo, R. (1957). Statistical-mechanical theory of irreversible processes. I. General theory and simple applications to magnetic and conduction problems. *J. Phys. Soc. Jpn.* 12: 570.
3 Craig, I.R. and Manolopoulos, D.E. (2004). Quantum statistics and classical mechanics: real time correlation functions from ring polymer molecular dynamics. *J. Chem. Phys.* 121: 3368–3373.
4 Tuckerman, M.E. (2002). Path integration via molecular dynamics. In: *Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms*, Lecture Notes, NIC Series, vol. 10 (eds. J. Grotendorst, D. Marx and A. Muramatsu), 269–298. Jülich: John von Neumann Institute for Computing.
5 Chandler, D. (1987). *Introduction to Modern Statistical Mechanics*. New York: Oxford University Press.
6 Mukamel, S. (1995). *Principles of Nonlinear Optical Spectroscopy*. Oxford University Press.
7 Egorov, S.A., Rabani, E., and Berne, B.J. (1998). Vibronic spectra in condensed matter: a comparison of exact quantum mechanical and various semiclassical treatments for harmonic baths. *J. Chem. Phys.* 108: 1407–1422.
8 Rabani, E., Egorov, S.A., and Berne, B.J. (1998). A comparison of exact quantum mechanical and various semiclassical treatments for the vibronic absorption spectrum: the case of fast vibrational relaxation. *J. Chem. Phys.* 109: 63766381.
9 Richardson, J.O., Meyer, P., Pleinert, M.-O., and Thoss, M. (2017). An analysis of nonadiabatic ring-polymer molecular dynamics and its application to vibronic spectra. *Chem. Phys.* 482: 124–134. arXiv:1609.00644 [physics.chem-ph].
10 Karsten, S., Ivanov, S.D., Bokarev, S.I., and Kühn, O. (2018). Quasi-classical approaches to vibronic spectra revisited. *J. Chem. Phys.* 148: 102337.
11 Miller, W.H., Schwartz, S.D., and Tromp, J.W. (1983). Quantum mechanical rate constants for bimolecular reactions. *J. Chem. Phys.* 79: 4889.
12 Wolynes, P.G. (1987). Imaginary time path integral Monte Carlo route to rate coefficients for nonadiabatic barrier crossing. *J. Chem. Phys.* 87: 6559.
13 Topaler, M. and Makri, N. (1996). Path integral calculation of quantum nonadiabatic rates in model condensed phase reactions. *J. Phys. Chem.* 100: 4430–4436.
14 Shi, Q. and Geva, E. (2004). Nonradiative electronic relaxation rate constants from approximations based on linearizing the path-integral forward-backward action. *J. Phys. Chem. A* 108: 6109–6116.
15 Huo, P., Miller, T.F. III, and Coker, D.F. (2013). Communication: predictive partial linearized path integral simulation of condensed phase electron transfer dynamics. *J. Chem. Phys.* 139: 151103.

16 Richardson, J.O. and Thoss, M. (2014). Non-oscillatory flux correlation functions for efficient nonadiabatic rate theory. *J. Chem. Phys.* 141: 074106. arXiv:1406.3144 [physics.chem-ph].

17 Kananenka, A.A., Sun, X., Schubert, A. et al. (2018). A comparative study of different methods for calculating electronic transition rates. *J. Chem. Phys.* 148: 102304.

18 Makri, N. (2015). Quantum-classical path integral: a rigorous approach to condensed phase dynamics. *Int. J. Quantum Chem.* 115: 1209–1214.

19 Meyer, H.-D. and Miller, W.H. (1979). A classical analog for electronic degrees of freedom in non-adiabatic collision processes. *J. Chem. Phys.* 70: 3214–3223.

20 Stock, G. and Thoss, M. (1997). Semiclassical description of nonadiabatic quantum dynamics. *Phys. Rev. Lett.* 78: 578–581.

21 Stock, G. and Thoss, M. (2005). Classical description of nonadiabatic quantum dynamics. *Adv. Chem. Phys.* 131: 243–376.

22 Tully, J.C. (1990). Molecular dynamics with electronic transitions. *J. Chem. Phys.* 93: 1061–1071.

23 Micha, D.A. (1983). A self-consistent eikonal treatment of electronic transitions in molecular collisions. *J. Chem. Phys.* 78: 7138–7145.

24 Stock, G. and Müller, U. (1999). Flow of zero-point energy and exploration of phase space in classical simulations of quantum relaxation dynamics. *J. Chem. Phys.* 111: 65–76.

25 Kelly, A., van Zon, R., Schofield, J., and Kapral, R. (2012). Mapping quantum- classical Liouville equation: projectors and trajectories. *J. Chem. Phys.* 136: 084101.

26 Runeson, J.E. and Richardson, J.O. (2019). Spin-mapping approach for nonadiabatic molecular dynamics. *J. Chem. Phys.* 151: 044119. arXiv:1904.08293 [physics.chem-ph]; Runeson, J.E. and Richardson, J.O. (2020). Generalized spin mapping for quantum-classical dynamics. *J. Chem. Phys.* 152: 084110.

27 Sun, X. and Miller, W.H. (1997). Semiclassical initial value representation for electronically nonadiabatic molecular dynamics. *J. Chem. Phys.* 106: 6346–6353.

28 Cotton, S.J., Liang, R., and Miller, W.H. (2017). On the adiabatic representation of Meyer-Miller electronic-nuclear dynamics. *J. Chem. Phys.* 147: 064112.

29 Mandal, A., Yamijala, S.S.R.K.C., and Huo, P. (2018). Quasi-diabatic representation for nonadiabatic dynamics propagation. *J. Chem. Theory Comput.* 14: 1828–1840.

30 Sun, X., Wang, H., and Miller, W.H. (1998). Semiclassical theory of electronically non-adiabatic dynamics: results of a linearized approximation to the initial value representation. *J. Chem. Phys.* 109: 7064–7074.

31 Wang, H., Song, X., Chandler, D., and Miller, W.H. (1999). Semiclassical study of electronically nonadiabatic dynamics in the condensed-phase: spin-boson problem with Debye spectral density. *J. Chem. Phys.* 110: 4828–4840.

32 Miller, W.H. (2009). Electronically nonadiabatic dynamics via semiclassical initial value methods. *J. Phys. Chem. A* 113: 1405–1415.

33 Kapral, R. (2006). Progress in the theory of mixed quantum-classical dynamics. *Annu. Rev. Phys. Chem.* 57: 129–157.

34 Kim, H., Nassimi, A., and Kapral, R. (2008). Quantum-classical Liouville dynamics in the mapping basis. *J. Chem. Phys.* 129: 084102.

35 Kapral, R. (2015). Quantum dynamics in open quantum-classical systems. *J. Phys.-Condens. Mat.* 27: 073201.

36 Church, M.S., Hele, T.J.H., Ezra, G.S., and Ananth, N. (2018). Nonadiabatic semiclassical dynamics in the mixed quantum-classical initial value representation. *J. Chem. Phys.* 148: 102326.

37 Bonella, S. and Coker, D.F. (2005). LAND-map, a linearized approach to nonadiabatic dynamics using the mapping formalism. *J. Chem. Phys.* 122: 194102.

38 Hsieh, C.-Y. and Kapral, R. (2012). Nonadiabatic dynamics in open quantum-classical systems: forward-backward trajectory solution. *J. Chem. Phys.* 137: 22A507.

39 Huo, P. and Coker, D.F. (2011). Communication: partial linearized density matrix dynamics for dissipative, non-adiabatic quantum evolution. *J. Chem. Phys.* 135: 201101.

40 Miller, W.H. and Cotton, S.J. (2016). Classical molecular dynamics simulation of electronically non-adiabatic processes. *Faraday Discuss.* 195: 9–30.

41 Weiss, U. (2012). *Quantum Dissipative Systems*, 4e. Singapore: World Scientific.

42 Garg, A., Onuchic, J.N., and Ambegaokar, V. (1985). Effect of friction on electron transfer in biomolecules. *J. Chem. Phys.* 83: 4491.

43 Marcus, R.A. (1993). Electron transfer reactions in chemistry. Theory and experiment. *Rev. Mod. Phys.* 65: 599.

44 Leggett, A.J., Chakravarty, S., Dorsey, A.T. et al. (1987). Dynamics of the dissipative two-state system. *Rev. Mod. Phys.* 59 (1).

45 Bader, J.S., Kuharski, R.A., and Chandler, D. (1990). Role of nuclear tunneling in aqueous ferrous- ferric electron transfer. *J. Chem. Phys.* 93: 230–236.

46 Mak, C.H. and Chandler, D. (1991). Coherent-incoherent transition and relaxation in condensed-phase tunneling systems. *Phys. Rev. A* 44: 2352–2369.

47 Makarov, D.E. and Makri, N. (1993). Tunneling dynamics in dissipative curve-crossing problems. *Phys. Rev. A* 48: 3626.

48 Thoss, M., Wang, H., and Miller, W.H. (2001). Self-consistent hybrid approach for complex systems: application to the spin-boson model with Debye spectral density. *J. Chem. Phys.* 115: 2991.

49 Kelly, A., Brackbill, N., and Markland, T.E. (2015). Accurate nonadiabatic quantum dynamics on the cheap: making the most of mean field theory with master equations. *J. Chem. Phys.* 142: 094110.

50 Montoya-Castillo, A. and Reichman, D.R. (2016). Approximate but accurate quantum dynamics from the Mori formalism: I. Nonequilibrium dynamics. *J. Chem. Phys.* 144: 184104.

51 Montoya-Castillo, A. and Reichman, D.R. (2017). Approximate but accurate quantum dynamics from the Mori formalism. II. Equilibrium time correlation functions. *J. Chem. Phys.* 146: 084110.

52 Müller, U. and Stock, G. (1999). Flow of zero-point energy and exploration of phase space in classical simulations of quantum relaxation dynamics. II. Application to nonadiabatic processes. *J. Chem. Phys.* 111: 77.

53 Wang, H., Thoss, M., and Miller, W.H. (2001). Systematic convergence in the dynamical hybrid approach for complex systems: a numerically exact methodology. *J. Chem. Phys.* 115: 2979.

54 Craig, I.R., Thoss, M., and Wang, H. (2007). Proton transfer reactions in model condensed-phase environments: accurate quantum dynamics using the multilayer multiconfiguration time-dependent Hartree approach. *J. Chem. Phys.* 127: 144503.

55 Craig, I.R. and Manolopoulos, D.E. (2005). Chemical reaction rates from ring polymer molecular dynamics. *J. Chem. Phys.* 122: 084106.

56 Kelly, A., Montoya-Castillo, A., Wang, L., and Markland, T.E. (2016). Generalized quantum master equations in and out of equilibrium: when can one win? *J. Chem. Phys.* 144: 184105.

57 Saller, M.A.C., Kelly, A., and Richardson, J.O. (2019). On the identity of the identity operator in nonadiabatic linearized semiclassical dynamics. *J. Chem. Phys.* 150: 071101.

58 Saller, M.A.C., Kelly, A., and Richardson, J.O. (2019). Improved population operators for multi-state nonadiabatic dynamics with the mixed quantum-classical mapping approach. *Faraday Discuss.*: 150–167. arXiv:1904.11847 [physics.chem-ph].

59 Chandler, D. and Wolynes, P.G. (1981). Exploiting the isomorphism between quantum theory and classical statistical mechanics of polyatomic fluids. *J. Chem. Phys.* 74: 4078–4095.

60 Parrinello, M. and Rahman, A. (1984). Study of an *F* center in molten KCl. *J. Chem. Phys.* 80: 860–867.

61 Ceriotti, M., Fang, W., Kusalik, P.G. et al. (2016). Nuclear quantum effects in water and aqueous systems: experiment, theory, and current challenges. *Chem. Rev.* 116: 7529–7550.

62 Markland, T.E. and Ceriotti, M. (2018). Nuclear quantum effects enter the mainstream. *Nat. Rev. Chem.* 2: 0109.

63 Habershon, S., Manolopoulos, D.E., Markland, T.E., and Miller, T.F. III, (2013). Ring-polymer molecular dynamics: quantum effects in chemical dynamics from classical trajectories in an extended phase space. *Annu. Rev. Phys. Chem.* 64: 387–413.

64 Richardson, J.O. and Thoss, M. (2013). Communication: nonadiabatic ring-polymer molecular dynamics. *J. Chem. Phys.* 139: 031102.

65 Ananth, N. (2013). Mapping variable ring polymer molecular dynamics: a path-integral based method for nonadiabatic processes. *J. Chem. Phys.* 139: 124102.

66 Chowdhury, S.N. and Huo, P. (2017). Coherent state mapping ring-polymer molecular dynamics for non-adiabatic quantum propagations. *J. Chem. Phys.* 147: 214109. arXiv:1706.08403 [physics.chem-ph].

67 Alexander, M.H. (2001). Path-integral simulation of finite-temperature properties of systems involving multiple, coupled electronic states. *Chem. Phys. Lett.* 347: 436–442.

68 Ceperley, D.M. (1995). Path integrals in the theory of condensed helium. *Rev. Mod. Phys.* 67: 279.

69 Zimmermann, T. and Vaníček, J. (2013). Role of sampling in evaluating classical time autocorrelation functions. *J. Chem. Phys.* 139: 104105.

70 Ananth, N. and Miller, T.F. III, (2010). Exact quantum statistics for electronically nonadiabatic systems using continuous path variables. *J. Chem. Phys.* 133: 234103.

71 Runeson, J., Nava, M., and Parrinello, M. (2018). Quantum symmetry from enhanced sampling methods. *Phys. Rev. Lett.* 121: 140602.

72 Kalos, M.H. and Whitlock, P.A. (1986). *Monte Carlo methods. Vol. 1: Basics*. New York, USA: Wiley-Interscience.

73 Light, J.C., Hamilton, I.P., and Lill, J.V. (1985). Generalized discrete variable approximation in quantum mechanics. *J. Chem. Phys.* 82: 1400.

74 Witt, A., Ivanov, S.D., Shiga, M. et al. (2009). On the applicability of centroid and ring polymer path integral molecular dynamics for vibrational spectroscopy. *J. Chem. Phys.* 130: 194510.

75 Rossi, M., Ceriotti, M., and Manolopoulos, D.E. (2014). How to remove the spurious resonances from ring polymer molecular dynamics. *J. Chem. Phys.* 140: 234166.

76 Müller, U. and Stock, G. (1998). Consistent treatment of quantum-mechanical and classical degrees of freedom in mixed quantum-classical simulations. *J. Chem. Phys.* 108: 7516–7526.

77 Bonella, S. and Coker, D.F. (2001). A semiclassical limit for the mapping Hamiltonian approach to electronically nonadiabatic dynamics. *J. Chem. Phys.* 114: 7778–7789.

78 Bonella, S. and Coker, D.F. (2001). Semi-classical implementation of mapping Hamiltonian methods for general non-adiabatic problems. *Chem. Phys.* 268: 189–200.

79 Richardson, J.O., Bauer, R., and Thoss, M. (2015). Semiclassical Green's functions and an instanton formulation of electron-transfer rates in the nonadiabatic limit. *J. Chem. Phys.*

143: 134115. arXiv:1508.04919 [physics.chem-ph]; Richardson, J.O. (2015). Ring-polymer instanton theory of electron transfer in the nonadiabatic limit. *J. Chem. Phys.* 143: 134116, arXiv:1508.05195 [physics.chem-ph].

80 Mattiat, J. and Richardson, J.O. (2018). Effects of tunnelling and asymmetry for system-bath models of electron transfer. *J. Chem. Phys.* 148: 102311. arXiv:1708.06702 [physics.chem-ph].

81 Heller, E.R. and Richardson, J.O. (2020). Instanton formulation of Fermi's golden rule in the Marcus inverted regime. *J. Chem. Phys.* 152: 034106.

82 Cao, J., Minichino, C., and Voth, G.A. (1995). The computation of electron transfer rates: the nonadiabatic instanton solution. *J. Chem. Phys.* 103: 1391.

83 Cao, J. and Voth, G.A. (1997). A unified framework for quantum activated rate processes. II. The nonadiabatic limit. *J. Chem. Phys.* 106: 1769.

84 Lawrence, J.E. and Manolopoulos, D.E. (2018). Analytic continuation of Wolynes theory into the Marcus inverted regime. *J. Chem. Phys.* 148: 102313.

85 Menzeleev, A.R., Bell, F., and Miller, T.F. III (2014). Kinetically constrained ring- polymer molecular dynamics for non-adiabatic chemical reactions. *J. Chem. Phys.* 140: 064103.

86 Duke, J.R. and Ananth, N. (2016). State space path integrals for electronically nonadiabatic reaction rates. *Faraday Discuss.* 195: 253–268.

Index

Quantum Chemistry and Dynamics of Excited States: Methods and Applications,
First Edition. Edited by Leticia González and Roland Lindh.

d

e

f

g

n

o

t