Title
Blasting off at the speed of light: Exploring relativistic quantum chemistry for core spectroscopy and heavy elements

Author
dos Anjos Cunha, Leonardo

Publication Date
2023

Peer reviewed|Thesis/dissertation

Blasting off at the speed of light: Exploring relativistic quantum chemistry for core spectroscopy and heavy elements

by

Leonardo dos Anjos Cunha

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Chemistry

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Martin Head-Gordon, Chair
Professor Eric Neuscamman
Professor Hartmut Häffner

Summer 2023

Blasting off at the speed of light: Exploring relativistic quantum chemistry for core spectroscopy and heavy elements

Abstract

Blasting off at the speed of light: Exploring relativistic quantum chemistry for core spectroscopy and heavy elements

by

Leonardo dos Anjos Cunha

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor Martin Head-Gordon, Chair

Quantum mechanics is one of the most predictive theories that we have at our disposal to explain natural phenomena with atomistic detail. Yet, applying it to complex system, such as chemical reactions, is an extremely demanding task. Approximations justified by physical principles are necessary to reduce the computational cost of these simulations. This dissertation analyzes two classes of problems in molecular quantum mechanics, highlighting the main ingredients necessary to qualitatively and quantitatively describe them. First, we look at how the choice of orbitals influences the description of strong-field ionization processes in simple systems, drawing an analogy to the well-known symmetry dilemma within quantum chemistry. We show that, through a simple model and a mean-field treatment, allowing the wavefunction to simply separate pairs of electrons is not enough to understand the ionization process, requiring additional flexibility to allow for spin rotations.

The second class of problems discussed in this dissertation concerns the description of excited states that arise when promoting one of the electrons in an inner-shell of a system. X-ray spectroscopy is a vibrant field that has been explored in great detail over recent years due to advances in light sources. From a theory perspective, these core-excited states also pose some challenges to existing quantum chemical methods, especially due to lack of relaxation effects after creating a hole in one of the core orbitals. We discuss different approaches to model X-ray emission and absorption, as well as the role of scalar relativistic effects in accurately modeling the core-excitation energies and the spectra of heavy elements. In doing so, we have devised both state-specific and linear-response methods to model core-excited states, expanding the toolbox available to quantum chemists to interpret new experiments in the X-ray range of the electromagnetic spectrum.

To my family, in particular Dad, Mom and Grandma.
For always making sure that I had the best opportunities to succeed.

Contents

Acknowledgments

No self is of itself alone. It has a long chain of intellectual ancestors. The "I" is chained to ancestry by many factors...This is not mere allegory, but an eternal memory.
Erwin Schrödinger

As I was finishing my masters in 2018, I had the opportunity to read Marie Curie's biography and came across a quote about how, "in science, we should be more interested in things than in persons". At the time, I did not really understand what one of the greatest scientific minds of the twentieth century was trying to convey. Five years later, I realize that without people, most scientific endeavors are doomed to fail. Humans are a social species and we rely on cooperation to succeed. This dissertation is, if anything, a collective effort by the many giants who have been kind enough to let me stand on their shoulders to see further and accomplish my goals. Certain and unfortunately, I will fail in providing a comprehensive list, but I wholeheartedly wish to thank everyone for their help over the years.

First, I would like to thank my advisors, Martin Head-Gordon and Bill McCurdy. Bill has shown me that there is much more to explore beyond the quantum mechanics of bound-states, and, even though I did not follow this fascinating field as much as he would like, I still learned a lot from him and his group. Martin helped me understand the intricacies of bound electronic structure. I cannot thank him enough for pushing me to keep going when I was not sure I could finish my PhD, and for giving me enough freedom to explore different corners of quantum chemistry. I will always rejoice our weekly meetings and how much I have learned from our scientific discussions. I would also like to thank Prof. Eric Neuscamman and Prof. Hartmut Häffner for taking the time to read this dissertation and making sure it was scientifically sound.

I should also thank my Brazilian advisors, Francisco, Luiz and Max. Francisco and Luiz gave me the opportunity to do research as an undergraduate student, and opened doors that I once thought to be unreachable. Max has been both an advisor and close friend over the years. His passion about science is contagious, and his tendency to mark all of our manuscripts in yellow helped me become a (hopefully) better scientist. I thank him for helping me decide where to go to pursue my PhD. His advice proved, as usual, to be the right answer.

I would like to thank the whole MHG group, including Drs. Joohno Lee, Luke Bertels, Srimukh Prasad, Katherine Oosterbaan, Matthias Loipersberger, Jonathan Wong, Elliot Rossomme, Josie Hendrix, Abdulrahman Aldossary, Adam Rettig, Alexandra McIsaac, James Shee, Justin Talbot, Kevin Carter-Fenk, Alistair Sterling, Avijit Shee, as well as Brad Ganoe, Jiashu Liang, Zhenling Wang, Hengyuan Shen, Xiao Liu, Wendy Billings, Devin Hernandez, Stephen Quiton, Zeyi Zhang and Tim Schramm. I have enjoyed spending time in the group area and discussing science (or any other random topic) with them, and I will dearly miss the coffee walks (or froyo walks, Sheng Kee walks...or simply random walks) at 2pm everyday. I would be negligent if I did not mention Dr. Diptarka Hait, who has been a source of inspiration (and, at times, frustration), as well as a great colleague over the years.

His knowledge of any aspect of quantum chemistry is commendable, and I hope that others have the privilege to learn as much as I did from him. I would also like to acknowledge Juan Arias-Martinez and Richard Kang for their friendship. Juan has been the best office mate I could have wished for, and I thank him for always lending me his ear when I wanted to complain about something not working. I hope our paths cross again in the future, hopefully in an office with more windows. I thank Richard for patiently listening to my crazy ideas about spin-orbit coupling. I should also thank Jessie Woodcock and Shobit Singh for always making sure I got paid in time and for solving any administrative issues. In particular, thanks Shobit, the MVP, for always bringing an enormous positive energy to the group after the pandemic and for being a good friend.

I would also like to acknowledge the impact that Dr. Rachel Clune, Dr. Tarini Hardikar, Scott Garner, Becky Hanscam, Connie Robinson, Oxana Andriuc, Leo Hamerlynck and Nugent Lew have had on my PhD journey. They are one of the reasons why I survived my first years in Berkeley, and I will miss our shenanigans, Friday night board games in the basement of Gilman Hall, and playing racket sports over the weekend. I would like to thank Rachel for walking home with me in the afternoons and always making sure that I was feeling more than just ok, and Scott for all the lovely days and random discussions over the years. Moreover, I should thank Tarini for always joining me in complaining about Q-Chem, and Becky for being the best volleyball partner. Finally, thanks to Sonja Tomasi and Harrison Tuckman for being amazing CHEM 221B students and always pushing me to improve as a teacher.

Outside the university, I should thank Panagiotis Zarkos and Subhayan Roychoudhury for being wonderful roommates. Panos has been a great friend, always providing a fresh perspective on life, and how to balance work and other activities. Shub has always been open to discuss science, and to help me navigate frustrations with academia. I would also like to thank Daniel Koda and Samuel Cerqueira for always letting me join them in hiking adventures, as well as Guilherme Costa and Lucas Morais for making sure I was doing well despite being a few thousand miles away.

My journey in Berkeley would not have been as fulfilling if I had not met Rebecca Tomann. Becca has energized the last year of my PhD, and I am grateful for her company and our trips to Yosemite, Point Lobos, Big Sur, San Francisco and beyond. I cannot wait to have more great experiences with her in the next chapter of our lives, hopefully with the same good food, laughter and joy that have been the defining marks of the past year.

Finally, none of this would have been possible without the incredible patience and superhuman strength of my family, especially my dad Vilmar, my mom Inês, and my grandma Neide. They likely will not understand much of this dissertation, but this is as much their accomplishment as it is mine. They have provided me with every opportunity to chase my own dreams, even if that meant sacrificing themselves or letting me move far away. I thank them for all their unwavering love and support. I am forever grateful for all that they have taught me over the years, and I hope that this dissertation helps me express a little bit of that gratitude towards them.

Chapter 1

Introduction

The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble
P. A. M. Dirac "Quantum mechanics of many–electron systems" *Proc. R. Soc. Lond. A*, **123**, 714–733, 1929.

At the turn of the twentieth century, most of the physical laws describing natural phenomena were believed to have been already discovered. Isaac Newton had already laid the foundations of what would be known as classical mechanics and universal gravitation, whereas James Clerk Maxwell had unified the description of electromagnetism. The so-called "ultraviolet catastrophe" [1] deeply changed this perspective and revolutionized science as a whole. By assuming that the electromagnetic radiation can be emitted or absorbed only in discrete, instead of continuous, packets of energy called quanta [2], Max Planck contributed to the advent of a new theory capable of explaining the microscopical underpinnings of the natural world.

This new subject, quantum mechanics (QM), is the foundation upon which most of modern physics and chemistry is built on. QM describes the interplay between matter and energy at the atomic level [3]. The very modern idea of the atom itself is, ultimately, a quantum mechanical concept [4], showcasing how important such theory is to the description of the most basic processes in chemistry. Despite its extraordinary ability to explain nature around us, QM suffers from a computational complexity problem, as alluded to by Paul Dirac in the quote that opens this chapter. Aside from a few small models that can be solved analytically, such as the hydrogen atom [5, 6], computing the quantum mechanical properties of complex molecular systems is undeniably an extremely demanding task. Exact numerical solutions are usually unfeasible for most systems of chemical interest. The field of quantum chemistry is therefore concerned in developing affordable approximations to extend the applicability and predictive power of QM to larger systems. The quantum chemist's toolbox comprises not only of physical ideas rooted in nature's reality, but also of an inherent

ability to implement complex mathematical equations as efficient and scalable computer programs.

In this chapter, we will briefly outline some of these tools, from the non-relativistic quantum theory for molecular systems and the most common approximations used to solve the time-independent Schrödinger equation, to an overview of how to reconcile the ideas from Einstein's theory of special relativity [7] and quantum mechanics. We will employ atomic units, *i.e.* define the mass of the electron (m_e), and the elementary charge (e) to be unity, along with the values for the reduced Planck's constant ($\hbar$) and the quantity $4\pi\varepsilon_0$, throughout this work, unless noted otherwise.

1.1 Overview of Non-Relativistic Quantum Chemistry

In classical mechanics, Newton's equation of motion (or its generalized version from Lagrange's or Hamilton's formalisms [8]) is the key ingredient to describe the dynamical properties of a system. In quantum mechanics, such a role is played by the time-dependent Schrödinger equation (Eq. 1.1)

$$\mathbf{H}\,|\Psi(t)\rangle = i\hbar\frac{\partial}{\partial t}\,|\Psi(t)\rangle \tag{1.1}$$

In Eq. 1.1, the dynamical variable, $|\Psi(t)\rangle$, encodes all of the information about the QM system and can be interpreted in a statistical sense. Once projected into real space (Eq. 1.2, assuming a single particle for simplicity, but without loss of generality), the quantity $|\Psi(\vec{x},t)|^2$ represents the probability distribution of the location of a particle at a given time.

$$\Psi(\vec{x},t) = \langle\vec{x}|\Psi(t)\rangle \tag{1.2}$$

The dynamics of $|\Psi(t)\rangle$ is dictated by the Hamiltonian operator $\mathbf{H}$. In the case where this Hamiltonian does not depend explicitly on time, separation of variables lead to the more common time-independent version of the Schrödinger equation (Eq. 1.3). This is usually the starting point for modeling within quantum chemistry, and solving Eq. 1.3 yields a complete basis of stationary states $\{|\psi_k\rangle\}$ and their respective energies $\{E_k\}$. The complexity problem of QM is hidden in the structure of the Hamiltonian operator and what types of interactions between particles it is able to describe.

$$\mathbf{H}\,|\psi_k\rangle = E_k\,|\psi_k\rangle \tag{1.3}$$

For chemical systems, composed of a set of nuclei and electrons, and in the absence of external electromagnetic fields, the (non-relativistic) Hamiltonian operator is given by:

$$\mathbf{H} = \mathbf{T}_e + \mathbf{T}_n + \mathbf{V}_{ne} + \mathbf{V}_{nn} + \mathbf{V}_{ee} \tag{1.4}$$

$\mathbf{T}_e$ and $\mathbf{T}_n$ are the kinetic energies of the electrons and nuclei respectively. $\mathbf{V}_{ne}$,$\mathbf{V}_{nn}$ and $\mathbf{V}_{ee}$ are purely electrostatic (Coulombic) terms that describe electron-nuclear attraction, nuclear-nuclear repulsion and electron-electron repulsion respectively. Assuming that the system is

composed of electrons labeled with indices $\{i, j \ldots\}$, and nuclei labeled $\{A, B \ldots\}$, we can express each one of the terms in the Hamiltonian as follows:

$$\mathbf{T}_n = \sum_A -\frac{1}{2M_A}\nabla_A^2 \tag{1.5}$$

$$\mathbf{T}_e = -\frac{1}{2}\sum_i \nabla_i^2 \tag{1.6}$$

$$\mathbf{V}_{ne} = -\sum_{i,A} \frac{Z_A}{\left|\vec{R}_A - \vec{r}_i\right|} \tag{1.7}$$

$$\mathbf{V}_{ee} = \frac{1}{2}\sum_{i,j;i\neq j} \frac{1}{|\vec{r}_i - \vec{r}_j|} \tag{1.8}$$

$$\mathbf{V_{nn}} = \frac{1}{2}\sum_{A,B;A\neq B} \frac{Z_A Z_B}{\left|\vec{R}_A - \vec{R}_B\right|} \tag{1.9}$$

$$\tag{1.10}$$

where $\{\vec{r}_i\}$ and $\{\vec{R}_A\}$ are the positions of the electrons and nuclei respectively, while $\{Z_A\}$ and $\{M_A\}$ are the nuclear charge and mass.

1.1.1 The Born-Oppenheimer Approximation

The task of solving the time-independent Schrödinger equation for the Hamiltonian in Eq. 1.4 is quite daunting. The stationary states (which we will interchangeably refer to as wavefunctions throughout this work) would not only depend on the positions of the many electrons of the system, but also on the position of the nuclei, leading to a high dimensional and complex problem. It is, however, worth noticing that nuclei are much heavier than electrons. Indeed, the lightest nuclei, a single proton in the hydrogen atom, is about three orders of magnitude more massive than a single electron. This leads to the first common approximation often used in quantum chemistry. The Born-Oppenheimer (BO) approximation [6, 9, 10] is based on the mismatch in the time scales of electronic and nuclear motions, which allows us to decouple the degrees of freedom of these two sets of particles. Within BO, one can write an electronic Hamiltonian $\mathbf{H_e}$ (Eq. 1.11) that describes the behavior of the electrons at a given set of fixed nuclear coordinates $\{\vec{R}\}$. Note that V_{nn} is now simply a scalar quantity, and does not affect the electronic wavefunction.

$$\mathbf{H}_e(\{\vec{R}\}) = \mathbf{T}_e + \mathbf{V}_{ne}(\{\vec{R}\}) + \mathbf{V}_{ee} + V_{nn}(\{\vec{R}\}) \tag{1.11}$$

While helpful in reducing the dimensionality of the problem at hand, employing the BO approximation comes with its consequences, such as failing to properly describe the properties of the system when the energy of two (or more) electronic states become near-degenerate at

a common set of nuclear coordinates [6, 11–13]. These cases, however, will not be considered in this work and we will assume BO is a valid approximation.

Is this first approximation enough to allow us to investigate complex chemical systems? Unfortunately, even with the purely electronic Hamiltonian, we would still have to deal with a high dimensional mathematical object (the electronic wavefunction) that depends on the coordinates of many electrons. The origin of the issue lies in the electron-electron repulsion term, $\mathbf{V}_{ee}$, in the Hamiltonian. $\mathbf{V}_{ee}$ is a two-body (or two-particle) operator that couples the motions of different electrons, hindering the separation of variables that would simplify the solution of the time-independent electronic Schrödinger equation.

1.1.2 Hydrogen-like Atoms

It is instructive, nonetheless, to look at the structure of such solutions when the electron-electron repulsion is not present in the Hamiltonian. One electron systems, such as the hydrogen atom or highly charged cations of other elements, provide us with a physically meaningful, yet simple and analytically solvable, example. These can potentially inspire other approximations to solve the electronic Hamiltonian for more complex multi-electronic systems. For these hydrogen-like atoms, the Hamiltonian reduces to

$$\mathbf{T}_e = -\frac{1}{2}\nabla^2; \mathbf{V}_{ne} = -\frac{Z}{r} \tag{1.12}$$

$$\mathbf{H}_e = \mathbf{T}_e + \mathbf{V}_{ne} \tag{1.13}$$

where the nucleus is placed at the origin, $\vec{r}$ is the position of the electron and Z is the nuclear charge.

The exact expression of the hydrogenic wavefunctions (henceforth referred to as orbitals, which will also be used to denote any one-electron wavefunction later on) are a bit convoluted and can be easily found in introductory quantum mechanics textbooks [5, 6]. However, it is worth noticing that, while the orbitals are characterized by a set of three quantum numbers $\{n, l, m\}$, each representing a different aspect of quantization, such as spatial orientation or number of radial and angular nodes of the orbitals, the energy levels themselves (E_{nlm}) reveal some degeneracies, as they only depend on the so-called principal quantum number n (Eq. 1.14). This is an important fact that will be explored later on this chapter, when relativistic effects are taken into account.

$$E_{nlm} = -\frac{Z^2}{2n^2} \tag{1.14}$$

1.1.3 Hartree-Fock theory: a bridge between hydrogen-like systems and multi-electronic molecules

1.1.3.1 The Slater Determinant as a Trial Wavefunction

One of the takeaways from the previous discussion is that, if the term in the Hamiltonian that couples the electronic motion ($\mathbf{V}_{ee}$) was to be neglected, we would potentially be able to solve

the electronic structure problem. In this case, even if we had a many-electron system, its wavefunction could be written as a product of one-electron functions (orbitals). This neglects the fact that electrons, as fermions, are indistinguishable from each other and have to obey Fermi-Dirac statistics [14, 15]. As a practical consequence, exchanging the position of any two electrons should result in an overall sign change in the total wavefunction, keeping all of the observables unchanged. Alternatively, the total electronic wavefunction of the system is said to be antisymmetric with respect to exchanging the coordinates of any two particles. Mathematically, this is expressed as

$$\Psi(\vec{r}_1, s_1; \vec{r}_2, s_2; \vec{r}_3, s_3 \ldots) = -\Psi(\vec{r}_2, s_2; \vec{r}_1, s_1; \vec{r}_3, s_3 \ldots) \tag{1.15}$$

where $\{\vec{r}_i, s_i\}$ represent the spatial and spin (or intrinsic angular momentum) coordinates of each electron. Moreover, if two electron had the same spatial and spin coordinates, Eq. 1.15 shows us that the total wavefunction of the system would vanish. This is known as the Pauli exclusion principle [16], and it is a consequence of requiring electrons to follow the Fermi-Dirac statistics.

The simplest way to enforce this antisymmetry requirement into the electronic wavefunction is to write it as a Slater determinant [17] whose rows and columns are composed of the single-particle orbitals $\{\chi_i, \chi_j \ldots \chi_N\}$ and the electronic (spatial and spin) coordinates:

$$\Phi(\vec{r}_1, s_1; \vec{r}_2, s_2; \vec{r}_3, s_3 \ldots) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_i(\vec{r}_1, s_1) & \chi_i(\vec{r}_2, s_2) & \chi_i(\vec{r}_3, s_3) \ldots \chi_i(\vec{r}_N, s_N) \\ \chi_j(\vec{r}_1, s_1) & \chi_j(\vec{r}_2, s_2) & \chi_j(\vec{r}_3, s_3) \ldots \chi_j(\vec{r}_N, s_N) \\ \vdots & \vdots & \ddots & \vdots \end{vmatrix} \tag{1.16}$$

where N is the total number of electrons in the system.

Instead of a single product of one-particle orbitals, we now have an antisymmetric linear combination of such products, enforcing the Fermi-Dirac statistics constraint into the electronic wavefunction. As we will discuss later on, using the Slater determinant as a trial wavefunction to solve the electronic structure problem has the physical consequence of simplifying the electron-electron repulsion term in the Hamiltonian. However, it is instructive to briefly discuss the properties of the single-particle orbitals that compose the Slater determinant.

1.1.3.2 On the Choice of Orbitals

So far, nothing has been said or assumed about the single-particle orbitals $\{\chi_i, \chi_j \ldots \chi_N\}$, aside from their dependence on the spatial and spin coordinates of a given electron. The eigenstates of the spin operator in any direction form a complete basis that can be used to provide us with more information about the spin component of these single-particle orbitals. The eigenstates $\{|\alpha\rangle, |\beta\rangle\}$ of the z component of the spin operator $\vec{\mathbf{S}}$ for an electron, a fermion with spin $s = \frac{1}{2}$, are given by [6]:

$$\mathbf{S}_z |\alpha\rangle = \frac{1}{2} |\alpha\rangle \quad \text{and} \quad \mathbf{S}_z |\beta\rangle = -\frac{1}{2} |\beta\rangle \tag{1.17}$$

The most general expression for a single-particle wavefunction is simply a linear combination of the z component eigenstates of the spin operator (more generally, any component could be used). The weights or coefficients of each eigenstate are given by the spatial components $\{\phi_i^\alpha(\vec{r}), \phi_i^\beta(\vec{r})\}$ of the orbital. This is expressed as follows:

$$\chi_i(\vec{r}, s) = \phi_i^\alpha(\vec{r})\alpha(s) + \phi_i^\beta(\vec{r})\beta(s) \tag{1.18}$$

This given rise to the generalized (G) class of orbitals. The spatial components $\{\phi_i^\alpha(\vec{r}), \phi_i^\beta(\vec{r})\}$ can be allowed to be complex if necessary, thus being named complex generalized (cG) orbitals. How should we go about reconciling this with our usual general chemistry knowledge, where orbitals are either α or β, but not a weird and complex combination of both? The answer lies in imposing constraints on the spatial and spin components of single-particle orbitals.

For instance, we can impose that we should have two separate sets, each with a given spin, either α or β, of orbitals composed of only real, but distinct, spatial parts (Eq. 1.19). This choice is usually referred to as unrestricted (U) orbitals. If one imposes the additional constraint that the spatial components of the α and β orbitals should be the same, this would lead to what is known as restricted (R) orbitals (Eq. 1.20), the most common choice for systems that do not have any unpaired electrons.

$$\chi_i(\vec{r}, s) = \phi_i^\alpha(\vec{r})\alpha(s) \quad \text{or} \quad \phi_i^\beta(\vec{r})\beta(s) \tag{1.19}$$

$$\chi_i(\vec{r}, s) = \phi_i(\vec{r})\alpha(s) \quad \text{or} \quad \phi_i(\vec{r})\beta(s) \tag{1.20}$$

Choosing a given class of orbitals (generalized, unrestricted, restricted) has its consequences. This is usually reflected on the symmetries obeyed by the electronic wavefunction. The electronic Hamiltonian in Eq. 1.11 commutes with several operators, such as the time-reversal $\mathbf{\Theta}$, complex conjugation $\mathbf{K}$, the total spin $\vec{\mathbf{S}}$ and $\mathbf{S^2}$ [18, 19]. The electronic wavefunction should, therefore, also be an eigenstate of all of these operators. For trial wavefunction composed of a single Slater determinant, this is only true when the choice the orbitals are real and restricted. For instance, if we allow the spatial parts of α and β orbitals to be different from one another, such as in the unrestricted case, the wavefunction is no longer an eigenfunction of $\mathbf{S^2}$ and $\mathbf{\Theta}$. Lifting all constraints and choosing cG orbitals leads to a Slater determinant that does not necessarily obey any of the symmetries of the Hamiltonian. A summary of the symmetries respected by each choice of orbitals is presented in Figure 1.1. These considerations will be further explored in the later chapters of this work. For now, we have the tools to discuss yet another approximation towards the solution of the electronic structure problem.

1.1.3.3 Variational Principle and the Hartree-Fock Method

So far, we have established the Slater determinant as our trial solution of the electronic Schrödinger equation, and that we have certain flexibility on the choice of the orbitals. We still need the tools, however, to obtain these orbitals, and finally calculate the electronic

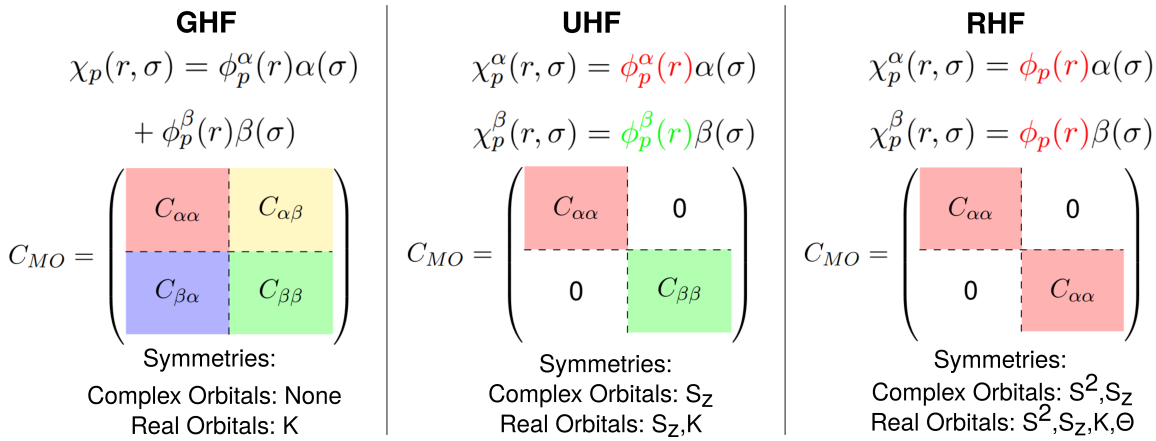


Figure 1.1: Different classes of orbitals, with structure of their respective coefficient matrix in a basis and the preserved symmetries of the exact Hamiltonian.

energy. We shall make use of the variational principle [5, 6], which states that the energy of any trial wavefunction is bounded by the true, exact ground-state energy of the system:

$$E_{exact} \leq \frac{\langle \Psi | \mathbf{H_{el}} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \tag{1.21}$$

Therefore, the best trial wavefunction, and orbitals, are the ones that minimize this expectation value. This is the Hartree-Fock (HF) method [20–23]: by using the variational principle, one can obtain the best set of single-particle orbitals that composes a Slater determinant and minimizes the expectation value of the electronic Hamiltonian. While the details of the full derivation can be found elsewhere [24, 25], here we emphasize a few important aspects. In evaluating the expectation value of the electronic Hamiltonian with respect to a Slater determinant we encounter the following terms:

$$E_0 = \langle \Phi | \mathbf{H}_e | \Phi \rangle = \langle \Phi | \mathbf{T}_e | \Phi \rangle + \langle \Phi | \mathbf{V}_{ne} | \Phi \rangle + \langle \Phi | \mathbf{V}_{ee} | \Phi \rangle + V_{nn} \tag{1.22}$$

$$V_{nn} = \frac{1}{2} \sum_{A,B;A \neq B} \frac{Z_A Z_B}{\left| \vec{R}_A - \vec{R}_B \right|} \tag{1.23}$$

$$\langle \Phi | \mathbf{T}_e | \Phi \rangle = -\frac{1}{2} \sum_i \int \chi_i^*(\vec{r}, s) \nabla^2 \chi_i(\vec{r}, s) d\vec{r} ds \tag{1.24}$$

$$\langle\Phi|\,\mathbf{V}_{ne}\,|\Phi\rangle = -\sum_i \int \chi_i^*(\vec{r},s)\sum_A \frac{Z_A}{\left|\vec{R}_A-\vec{r}\right|}\chi_i(\vec{r},s)d\vec{r}ds \tag{1.25}$$

$$\langle\Phi|\,\mathbf{V}_{ee}\,|\Phi\rangle = \frac{1}{2}\sum_{i,j;i\neq j}(J_{ij}-K_{ij}) \tag{1.26}$$

$$J_{ij} = \int \chi_i^*(\vec{r}_1,s_1)\chi_j^*(\vec{r}_2,s_2)\frac{1}{|\vec{r}_1-\vec{r}_2|}\chi_i(\vec{r}_1,s_1)\chi_j(\vec{r}_2,s_2)d\vec{r}_1ds_1d\vec{r}_2ds_2 \tag{1.27}$$

$$K_{ij} = \int \chi_i^*(\vec{r}_1,s_1)\chi_j^*(\vec{r}_2,s_2)\frac{1}{|\vec{r}_1-\vec{r}_2|}\chi_j(\vec{r}_1,s_1)\chi_i(\vec{r}_2,s_2)d\vec{r}_1ds_1d\vec{r}_2ds_2 \tag{1.28}$$

$\langle\Phi|\,\mathbf{T}_e\,|\Phi\rangle$ and $\langle\Phi|\,\mathbf{V}_{ne}\,|\Phi\rangle$ are sums over the expectation value of these one-electron operators for each occupied orbital.

The most interesting term is, as discussed earlier, the expectation value of $\langle\Phi|\,\mathbf{V}_{ee}\,|\Phi\rangle$, as it involves a two-particle operator. Using a Slater determinant as a trial wavefunction allows us to rewrite $\langle\Phi|\,\mathbf{V}_{ee}\,|\Phi\rangle$ as a sum of two components: J_{ij} and K_{ij}. The former represents the classical Coulomb repulsion between the probability distributions/densities of each orbital. The latter is a consequence of imposing Fermi-Dirac statistics in the wavefunction, thus being a purely quantum mechanical effect and being denoted as the exchange term/contribution. It is worth noticing that the exchange term only acts between pairs of orbitals that have the same spin (or components of the orbitals in the cG case).

By minimizing the energy expectation value (Eq. 1.22) with respect to the single-particle orbitals, we obtain the so called HF equations (Eq. 1.29), an eigenvalue problem that, once solved, yields the best orbitals that compose a single Slater determinant wavefunction.

$$\mathbf{f}\,|\chi_i\rangle = \varepsilon_i\,|\chi_i\rangle \tag{1.29}$$

$$\mathbf{f} = \mathbf{h} + \sum_j \mathbf{J}_j - \mathbf{K}_j \tag{1.30}$$

$$\mathbf{h} = -\frac{1}{2}\nabla^2 - \sum_A \frac{Z_A}{\left|\vec{R}_A-\vec{r}\right|} \tag{1.31}$$

$$\mathbf{J}_j\chi_i(\vec{r},s) = \left(\int \frac{|\chi_j(\vec{r}_2,s_2)|^2}{|\vec{r}-\vec{r}_2|}d\vec{r}_2ds_2\right)\chi_i(\vec{r},s) \tag{1.32}$$

$$\mathbf{K}_j\chi_i(\vec{r},s) = \left(\int \frac{\chi_j^*(\vec{r}_2,s_2)\chi_i(\vec{r}_2,s_2)}{|\vec{r}-\vec{r}_2|}d\vec{r}_2ds_2\right)\chi_j(\vec{r},s) \tag{1.33}$$

where ε_i is a Lagrange multiplier resulting from requiring the HF orbitals to remain orthonormal. It is possible to attach physical meaning to these quantities through the so-called Koopman's theorem, which states that the negative of ε_i correspond the ionization energy of an electron occupying χ_i, assuming no relaxation upon the removal of the electron [26]. An interesting feature of the HF equation is that, in order to obtain the set of orbitals

$\{\chi_i, \chi_j \ldots\}$ by solving Eq. 1.29, we first need to guess what this set looks like, as they are key ingredients to compute the Coulomb (Eq. 1.32) and exchange (Eq. 1.33) contributions to the Fock operator. This is then accomplished through an iterative protocol, or self-consistent field scheme, that is discussed in greater detail in Ref. 24.

1.1.3.4 Mean-Field Hamiltonian and Correlation Energy

Finally, Eqs. 1.29, 1.32 and 1.33 remind us of our initial goal: find an approximation that would allow us to treat the electronic Hamiltonian as a sum of non-interacting single-particle terms. Consequently, we would then be able to write the wavefunction of the system as an antisymmetric combination of products of one-electron functions. The HF method finally allows us to accomplish this goal: by using a single Slater determinant as our total wavefunction, we can write an effective one-electron equation that gives us the single-particle orbitals. Moreover, it is worth pointing out the physical content hidden in Eqs. 1.32 and 1.33. They indicate that, in the HF method, each electron experiences effectively an average field generated by the other electrons. This is of paramount importance to understand the limitations of HF theory and to improve upon it. We can define a many-body operator $\mathbf{F}$ that is composed by the sum of single-particle Fock operators $\mathbf{f}$ (Eq. 1.29):

$$\mathbf{F} = \sum_i \mathbf{f}(\vec{r_i}, s_i) \tag{1.34}$$

Due to the mean-field nature of the electron-electron interaction in HF, its approximate energy is not equal to the exact value that one would obtain from solving the full electronic Hamiltonian (Eq. 1.11). This allows us to define the correlation energy [27] as follows:

$$E_{\text{corr}} = E_{\text{exact}} - E_{\text{HF}} \tag{1.35}$$

Correlation effects can then be recovered by including certain contributions of the so-called fluctuation potential in post HF methods, as we will discuss in the following section.

$$\mathbf{V} = \mathbf{H}_e - \mathbf{F} \tag{1.36}$$

1.1.4 Climbing up the Wavefunction Mountain of Correlated Methods

HF is the simplest electronic structure method, as it builds upon the concept of trying to approximate the full electronic Hamiltonian through a sum of independent one-electron operators. This task is accomplished by using a single Slater determinant. HF does not yield the exact energy of a system. Indeed, a good estimate is that the correlation energy accounts for about 1 eV per electron pair [24]. While this does not seem to be much, it is critical to accurately model chemical process. These usually involve breaking and forming of chemical bonds (and likely changing the number of electron pairs in the system) and rely on

relative energies between reactants and products. For instance, an analysis of the simple case of the H_2 molecule, which has a single electron pair, reveals that the estimate for missing correlation energy corresponds to about 100 kJ/mol, which is on the same order of magnitude as the bond dissociation energy of the system. It is necessary to include correlation effects on top of the HF energy and its wavefunction in order to obtain an accurate description of physics and chemistry. This can be done through a myriad of different methods, of which we highlight only a few in this work.

1.1.4.1 Møller-Plesset Perturbation Theory

In Quantum Mechanics, Rayleigh-Schrödinger (RS) perturbation theory [28] is usually the most straightforward way to account for effects that were not previously included in the zeroth order description of a problem. This is accomplished by partitioning the total Hamiltonian into two pieces:

$$\mathbf{H} = \mathbf{H}_0 + \lambda \mathbf{V} \tag{1.37}$$

where $\mathbf{H}_0$ and $\mathbf{V}$ are, respectively, the easily solvable part of the full Hamiltonian (*i.e.* zeroth-order Hamiltonian) and the perturbation that we, in principle, do not know how to account for. λ is a simple parameter that controls the strength of the perturbation, but it is usually just used as a derivation tool. The same partitioning scheme can be applied to the total wavefunction and energy of the system:

$$|\Psi\rangle = \left|\Psi^{(0)}\right\rangle + \lambda \left|\Psi^{(1)}\right\rangle + \lambda^2 \left|\Psi^{(2)}\right\rangle + \dots \tag{1.38}$$

$$E = E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + \dots \tag{1.39}$$

where $\{E^{(0)}, E^{(1)} \dots\}$ and $\{\left|\Psi^{(0)}\right\rangle, \left|\Psi^{(1)}\right\rangle \dots\}$ are the corrections to the energy and wavefunction at each order of the perturbation, respectively.

By substituting Eqs. 1.37, 1.38 and 1.39 into Eq. 1.3 and collecting the terms by orders of λ, we obtain (up to second order):

$$\mathbf{H}_0 \left|\Psi^{(0)}\right\rangle = E^{(0)} \left|\Psi^{(0)}\right\rangle \tag{1.40}$$

$$\mathbf{H}_0 \left|\Psi^{(1)}\right\rangle + \mathbf{V} \left|\Psi^{(0)}\right\rangle = E^{(0)} \left|\Psi^{(1}\right\rangle + E^{(1)} \left|\Psi^{(0)}\right\rangle \tag{1.41}$$

$$\mathbf{H}_0 \left|\Psi^{(2)}\right\rangle + \mathbf{V} \left|\Psi^{(1)}\right\rangle = E^{(0)} \left|\Psi^{(2}\right\rangle + E^{(1)} \left|\Psi^{(1)}\right\rangle + E^{(2)} \left|\Psi^{(0)}\right\rangle \tag{1.42}$$

After some manipulation, we can obtain the expressions for the first and second order corrections to the energy (we will omit the expressions for the wavefunction corrections for simplicity):

$$E^{(1)} = \left\langle \Psi^{(0)} \left| \mathbf{V} \right| \Psi^{(0)} \right\rangle \tag{1.43}$$

$$E^{(2)} = \sum_{n \neq 0} \frac{\left| \left\langle \Psi_0^{(0)} \left| \mathbf{V} \right| \Psi_n^{(0)} \right\rangle \right|^2}{E_0^{(0)} - E_n^{(0)}} \tag{1.44}$$

where $E_n^{(0)}$ and $\left|\Psi_n^{(0)}\right\rangle$ are, respectively, the energy and wavefunction of the n-th eigenstate of the zeroth order Hamiltonian.

We are now in position to apply RS perturbation theory to the electronic Hamiltonian, which will be partitioned in a zeroth order part composed by the mean-field Hamiltonian (Eq. 1.34) and the fluctuation potential as the perturbation. This is the Møller-Plesset Perturbation Theory (MP) [29, 30]. In this case, we should note that the zeroth order excited states correspond to Slater determinants that can be obtained by including orbitals that were previously unoccupied in the ground-state Slater determinant. These orbitals will be label by $\{a, b, c \ldots\}$, whereas the occupied ground-state orbitals will be identified by $\{i, j, k \ldots\}$. We should also introduce the following notation to label these excited Slater determinants:

$$\left|\Psi_i^a\right\rangle = \mathbf{a}_a^\dagger \mathbf{a}_i \left|\Psi_0\right\rangle \tag{1.45}$$

$$\left|\Psi_{ij}^{ab}\right\rangle = \mathbf{a}_a^\dagger \mathbf{a}_b^\dagger \mathbf{a}_j \mathbf{a}_i \left|\Psi_0\right\rangle \tag{1.46}$$

where $\mathbf{a}_p^\dagger$ and $\mathbf{a}_p$ are the usual fermionic creation and annihilation operators for an electron at orbital p [25].

With these, the MP energy corrections through second order are given as:

$$E^{(0)} = \sum_i \varepsilon_i \tag{1.47}$$

$$E^{(1)} = E_{\mathrm{HF}} - \sum_i \varepsilon_i \tag{1.48}$$

$$E^{(2)} = -\frac{1}{4} \sum_{i \neq j, a \neq b} \frac{\left|\langle ij||ab\rangle\right|^2}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j} \tag{1.49}$$

where $\langle ij||ab\rangle$ is defined as follows:

$$\begin{aligned} \langle ij||ab\rangle = &\int \chi_i^*(\vec{r}_1, s_1)\chi_j^*(\vec{r}_2, s_2)\frac{1}{|\vec{r_1} - \vec{r_2}|}\chi_a(\vec{r}_1, s_1)\chi_b(\vec{r}_2, s_2)d\vec{r_1}ds_1 d\vec{r_2}ds_2 \\ &- \int \chi_i^*(\vec{r}_1, s_1)\chi_j^*(\vec{r}_2, s_2)\frac{1}{|\vec{r_1} - \vec{r_2}|}\chi_b(\vec{r}_1, s_1)\chi_a(\vec{r}_2, s_2)d\vec{r_1}ds_1 d\vec{r_2}ds_2 \end{aligned} \tag{1.50}$$

The first order MP correction to the energy does not improve over Hartree-Fock. In fact, we only recover the HF energy at first order. Hence, we only start recovering the correlation energy at second order (MP2) and Eq. 1.49 indicates the primary role played by doubly substituted determinants, $\left|\Psi_{ij}^{ab}\right\rangle$, in describing correlation effects.

While being a widely used tool to partially recover electronic correlation, MP theory has a few issues. One of them is the seemingly arbitrariness in partitioning the electronic Hamiltonian. Indeed, other perturbative approaches have been devised, with different partitioning schemes that yield potentially different results [31–35]. Other issues include lack

of guaranteed convergence of the perturbative series depending on the system under investigation, and the fact that the MP energies and wavefunctions are not variational. This not only means that there is no lower bound for the MP energy, but also calculating properties, such as dipoles moments and nuclear gradients, are more intricate tasks than in variational theories, such as HF [30, 36].

1.1.4.2 Configuration Interaction: adding more Slater determinants to the wavefunction

Despite these issues, MP theory reveals an appealing strategy to recover the correlation energy: adding more Slater determinants, through substitutions (or excitations) of the originally occupied orbitals, in our wavefunction. This is accomplished through configuration interaction (CI) theory [37, 38]. In CI theory, a linear combination of Slater determinants is used as the trial wavefunction. Assuming intermediate normalization, that is $\langle \Psi_{CI} | \Psi_0 \rangle = 1$, we have:

$$|\Psi_{CI}\rangle = |\Psi_0\rangle + \sum_{i,a} c_i^a \left|\Psi_i^a\right\rangle + \frac{1}{4} \sum_{i \neq j, a \neq b} c_{ij}^{ab} \left|\Psi_{ij}^{ab}\right\rangle + \ldots \tag{1.51}$$

where $|\Psi_0\rangle$ is the HF determinant, and $\{c_i^a, c_{ij}^{ab}, \ldots\}$ are the amplitudes, or CI coefficients, of the excited determinants $\{|\Psi_i^a\rangle, \left|\Psi_{ij}^{ab}\right\rangle, \ldots\}$. We can use the variational principle once again, but this time minimize the energy expectation value with respect to the CI coefficients, which yields the following eigenvalue equation:

$$\begin{bmatrix} E_0 & 0 & H_{0D} & 0 & 0 & \cdots \\ 0 & H_{SS} & H_{SD} & H_{ST} & 0 & \cdots \\ H_{0D} & H_{SD} & H_{DD} & H_{DT} & H_{DQ} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} 1 \\ C_S \\ C_D \\ \vdots \end{bmatrix} = E \begin{bmatrix} 1 \\ C_S \\ C_D \\ \vdots \end{bmatrix} \tag{1.52}$$

where $\{C_S, C_D, \ldots\}$ are the coefficients of the singly and doubly substituted, etc Slater determinants included in the expansion of the wavefunction. Moreover, the blocks of matrix elements $\{H_{OD}, H_{SS}, H_{SD}, \ldots\}$ are calculated between different Slater determinants:

$$H_{0D} = \left\langle \Psi_0 \left| \mathbf{H}_e \right| \Psi_{ij}^{ab} \right\rangle \tag{1.53}$$

$$H_{SS} = \left\langle \Psi_i^a \left| \mathbf{H}_e \right| \Psi_j^b \right\rangle \tag{1.54}$$

$$H_{SD} = \left\langle \Psi_i^a \left| \mathbf{H}_e \right| \Psi_{ij}^{ab} \right\rangle \tag{1.55}$$

We note that the electronic Hamiltonian, being a two-body operator, can not connect Slater determinants that differ from each other by more than two orbitals. This is the reason why $H_{0T} = H_{0Q} = \ldots = 0$. Moreover, the Brillouin's Theorem and the structure of the Hartree-Fock wavefunction, in its canonical form, ensures that $H_{0S} = 0$ [24]. This allows

us to write an exact expression of the correlation energy in terms of the coefficients of the doubly excited determinants:

$$E_{\text{corr}} = \frac{1}{4} \sum_{i \neq j, a \neq b} c_{ij}^{ab} \left\langle \Psi_0 \left| \mathbf{H}_e \right| \Psi_{ij}^{ab} \right\rangle \tag{1.56}$$

Once again, the importance of including double excitations on top of the HF wavefunction is highlighted. A note of caution is necessary however: this observation does not mean that other determinants are not important in the expansion of the wavefunction. As can be seen from the structure of Eq. 1.52, the double excitations are coupled to other ones, such as single, triple and quadruple excitations. For instance, projecting the CI Schrödinger equation by $\left\langle \Psi_l^d \right|$ allows us to (partially) see this coupling more clearly:

$$\sum_{i,a} c_i^a \left\langle \Psi_l^d \left| \mathbf{H}_e \right| \Psi_i^a \right\rangle + \sum_{i<j, a<b} c_{ij}^{ab} \left\langle \Psi_l^d \left| \mathbf{H}_e \right| \Psi_{ij}^{ab} \right\rangle + \sum_{i<j, a<b} c_{lij}^{dab} \left\langle \Psi_l^d \left| \mathbf{H}_e \right| \Psi_{lij}^{dab} \right\rangle = E c_l^d \tag{1.57}$$

This plays an important role in determining the coefficients of the doubly excited determinants, as the coupling leads to a system of equations that need to be solved simultaneously. Hence, the coefficients for the doubly excited determinants are modulated through their coupling to other excitations (Eq. 1.57), which renders the correlation energy in Eq. 1.56 also implicitly dependent on these other excitations. Finally, we note that Eq. 1.56 does not hold if $H_{0S} \neq 0$ (*e.g.* when non-HF references are used). In this case, the correlation energy depends explicitly on the coefficients for singly excited determinants as well.

While CI theory can, in principle, be made exact by including all possible excited Slater determinants (within a finite basis), it is not tractable, as the cost of evaluating the matrix elements and storing some of the computable intermediates is prohibitive for most complex systems, even when highly specialized algorithms are used [39–42]. An obvious alternative would be to truncate the CI expansion, leading to the hierarchy of CI methods that only contain single substitutions (CIS), singles and doubles (CISD) and so on.

CIS does not provide any improvement for the ground-state wavefunction and properties, as the singly excited determinants do not couple with the HF reference. This property, however, allows CIS to yield variational approximations to excited states that are orthogonal to the HF ground state. On the other hand, despite seemingly offering a viable path to recover the ground-state correlation energy, CISD and other truncated versions of CI theory fail in delivering consistent results for systems of increasing complexity/size. A simple, yet instructive way, of illustrating this issue is the analysis of the non-interacting H_2 dimer. Each monomer contains only two electrons, hence a truncation at the CISD level would be enough to exactly describe their wavefunction. The dimer, however, contains four electrons and CISD would not yield the exact energy of the system, and we would need to include up to quadruple excitations (CISDTQ) in the wavefunction to recover the exact correlation energy of the dimer.

This limitation is known as lack of size consistency [25]. Size consistency is the ability to recover product separable wavefunctions and additive energies when two fragments of a

system are not interacting. Due to the linear ansatz for the wavefunction, truncated CI theory does not include enough flexibility to properly describe the total wavefunction of the system. In our example, this translates to not allowing the total wavefunction to separate into products of monomer wavefunctions in the non-interacting limit.

1.1.4.3 Exponential Ansatz of the Wavefunction and Coupled-Cluster Theory

CI theory is based on a linear combination of substituted Slater determinants, and truncation of the CI expansion leads to lack of size consistency. An alternative way of consistently recovering the correlation energy is to use an exponential ansatz for the wavefunction as follows:

$$|\Psi_{CC}\rangle = e^{\mathbf{T}} |\Psi_0\rangle \tag{1.58}$$

$$\mathbf{T} = \mathbf{T}_1 + \mathbf{T}_2 + \mathbf{T}_3 + \dots \tag{1.59}$$

$$\mathbf{T}_1 |\Psi_0\rangle = \sum_{i,a} t_i^a |\Psi_i^a\rangle \tag{1.60}$$

$$\mathbf{T}_2 |\Psi_0\rangle = \frac{1}{4} \sum_{i \neq j, a \neq b} t_{ij}^{ab} |\Psi_{ij}^{ab}\rangle \tag{1.61}$$

where $\{t_\mu\}$ is an amplitude/weight associated with a given excited Slater determinant $|\Psi_\mu\rangle$. This is the couple-cluster (CC) wavefunction and the foundation of CC theory [43–46].

Without truncation, CC is equivalent to CI theory. This can be seen by expanding the exponential cluster operator and grouping terms by order of excitation:

$$e^{\mathbf{T}} = \mathbf{I} + \mathbf{T} + \frac{1}{2!}\mathbf{T}^{\mathbf{2}} + \dots \tag{1.62}$$

$$= \mathbf{I} + \mathbf{T}_1 + \left(\mathbf{T}_2 + \frac{1}{2!}\mathbf{T}_1^2\right) + \dots \tag{1.63}$$

All single excitations are generated by the $\mathbf{T}_1$ operator, while all doubles excitations are generated by the $\mathbf{T}_2$ and $\mathbf{T}_1^2$ operators. Truncated CC theory is more useful than CI as higher orders of excitations can be included even if the cluster operator $\mathbf{T}$ is truncated at lower order. For instance, CCSD, with $\mathbf{T} = \mathbf{T}_1 + \mathbf{T}_2$, is able to include some quadruple excitations in the wavefunction, a feature that can not be achieved at the same level of truncation in CI. The exponential ansatz also yields CC energies that are size consistent, as it allows for product separability of non-interacting fragment wavefunctions [25, 46].

With all of these nice properties, are there any disadvantages in using CC theory to solve the electronic structure problem? Unfortunately, obtaining the CC energies and cluster amplitudes is not as straightforward as in CI theory. If we were to calculate the expectation value of the electronic Hamiltonian with respect to the CC wavefunction (even a truncated version of it), the expansion would be as computationally demanding as full configuration interaction (FCI). This renders variational CC [47, 48] impractical for any complex system.

The alternative approach is based on a projection scheme that makes use of the Baker-Haussdorf-Campbell (BCH) expansion:

$$e^{-\mathbf{A}}\mathbf{B}e^{\mathbf{A}} = \mathbf{B} + [\mathbf{B}, \mathbf{A}] + \frac{1}{2!}[[\mathbf{B}, \mathbf{A}], \mathbf{A}] + \dots \tag{1.64}$$

where $\mathbf{A}$ and $\mathbf{B}$ are arbitrary operators. In the case of the electronic Hamiltonian, which is a two body operator, in the context of CC theory and the cluster operator $\mathbf{T}$, the BCH expansion truncates after four nested commutators:

$$\bar{H} = e^{-\mathbf{T}}\mathbf{H}e^{\mathbf{T}} = \mathbf{H} + [\mathbf{H}, \mathbf{T}] + \frac{1}{2!}[[\mathbf{H}, \mathbf{T}], \mathbf{T}] + \frac{1}{3!}[[[\mathbf{H}, \mathbf{T}], \mathbf{T}], \mathbf{T}] + \frac{1}{4!}[[[[\mathbf{H}, \mathbf{T}], \mathbf{T}], \mathbf{T}], \mathbf{T}] \tag{1.65}$$

Within this formalism, the projected CC equations yield expressions for the CC ground-state energy and the cluster amplitudes within the truncated subspace. For instance, in CCSD, the cluster operator is truncated at the singles and doubles excitation, and the corresponding subspace then includes the HF reference determinant, and its singly and doubly excited counterparts. Projecting the CC Schrödinger equation (Eq. 1.66) into this subspace gives us a system of equations to obtain the CC energy and amplitudes.

$$\bar{H}\Psi_0 = E_{\mathrm{CC}}\Psi_0 \tag{1.66}$$

$$E_{\mathrm{CC}} = \left\langle \Psi_0 \middle| \bar{H} \middle| \Psi_0 \right\rangle \tag{1.67}$$

$$\left\langle \Psi_i^a \middle| \bar{H} \middle| \Psi_0 \right\rangle = 0 \tag{1.68}$$

$$\left\langle \Psi_{ij}^{ab} \middle| \bar{H} \middle| \Psi_0 \right\rangle = 0 \tag{1.69}$$

MP and CC theories are intrinsically related to each other. Similar to MP theory, calculation of properties within the CC formalism is not a simple task, due to the its non-variational character. Moreover, both MP and CC theory (especially low orders of truncation) rely, to some extent, on the HF wavefunction being a qualitative good starting point. In other words, they require that there is a single dominant Slater determinant in the electronic wavefunction. This configuration is hopefully captured by the HF reference, while the add-on excited configurations (either through perturbation theory or the exponential cluster operator) contribute only slightly to the total wavefunction and energy. This leads to a (somewhat arbitrary) distinction between strong (or static) and weak (or dynamical) correlation. MP and CC are effective methods to recover weak correlation. The problem of recovering strong correlation is out of the scope of this work and we refer the interested reader to additional references about the topic [49, 50].

1.1.5 DFT and the Central Role of the Electronic Density

So far, all of the electronic structure methods presented rely on having the wavefunction of the system play a central role. Indeed, the Schrödinger equation, either in its time-dependent or

time-independent version, gives us a recipe for how to obtain and manipulate such quantities. One of the disadvantages of such approach is the complexity of these mathematical objects. The electronic wavefunction for a system containing N particles will depend on $4N$ variables, three spatial and a spin coordinate for each electron. Recall that our main goal in the beginning of this chapter was to list the available tools quantum chemists have at their disposal to reduce the complexity of the often daunting task of exactly solving the electronic structure problem. Can we do better than simply relying on this often impractically large Hilbert space?

We have learned along the way some nice properties that we would like to include in an "ideal" theory that depends on less information than the $4N$ variables of the wavefunction space. For instance, we would like to keep the simplicity of HF theory in dealing with the electron-electron repulsion term of the Hamiltonian, while also being able to recover correlation effects in order to model chemical processes. This is where density functional theory (DFT), and more specifically Kohn-Sham (KS) DFT comes to our rescue. KS-DFT emphasizes the leading role of the charge (or number) density, a much more amenable mathematical object that depends only on three spatial coordinates regardless of the number of electrons in the system, to solve the electronic structure problem.

1.1.5.1 Hohenberg-Kohn Theorems as the Foundation of DFT

While the idea of using the charge density as the primary ingredient to determine the electronic properties of a system was pioneered by Thomas, Fermi and Dirac [51–53] in the early days of quantum mechanics, their model failed to predict chemical bonding [54], and therefore did not encounter widespread use among quantum chemists. The framework, however, was greatly expanded decades later by Hohenberg and Kohn (HK), laying the main foundations of modern DFT [55].

Modern DFT is rooted in two theorems. For a system of interacting electrons under the influence of an external one-body local potential $v_{ext}(\vec{r})$

1. There is a 1:1 mapping between the system's charge density $\rho(\vec{r})$ and the external potential $v_{ext}(\vec{r})$ (aside from an additive constant) that it is subjected to. Once $v_{ext}(\vec{r})$ is known, the whole Hamiltonian of the system is completely specified and the energies of its eigenstates can be calculated exactly. Consequently, there is an energy functional $E[\rho(\vec{r})]$ that maps the ground-state density of the system into its energy.

2. Let $\rho_0(\vec{r})$ be the charge density of the ground-state of the system. For any arbitrary (v-representable, that is originating from a proper electronic wavefunction for a valid $v_{ext}(\vec{r})$) density $\rho(\vec{r})$, we have $E[\rho_0(\vec{r})] \leq E[\rho(\vec{r})]$. This is the equivalent of the variational principle in DFT.

In our case, the external potential is, in the absence of other fields, the electron-nuclear attraction, $\mathbf{V_{ne}}$. Therefore, we will replace that term in the electronic Hamiltonian by $v_{ext}(\vec{r})$. Despite their importance in laying the groundwork for modern DFT, the HK theorems do not

necessarily point to a practical path ahead on how to actually obtain the exact ground-state density, or even calculate the energy once said density is available. The second HK theorem also does not provide any useful criteria to check if a trial density is in fact v-representable, yielding a trial space too large for the minimization process.

Fortunately, Levy proposed a constrained minimization protocol [56] that replaced the v-representability requirement by a somewhat looser criteria known as the n-representability condition. In this case, the Hilbert space is broken down into small sectors, each corresponding to a set of wavefunctions $\{|\Psi_\rho\rangle\}$ that are constrained to have the same density $\rho(\vec{r})$. Mathematically, this can be expressed as:

$$E[\rho_0(\vec{r})] = \langle \Psi[\rho_0(\vec{r})] | \mathbf{T_e} + \mathbf{V_{ee}} + \mathbf{v_{ext}}(\vec{r}) | \Psi[\rho_0(\vec{r})] \rangle \tag{1.70}$$

$$E[\rho_0(\vec{r})] = \min_{\rho} \left(\min_{|\Psi_\rho\rangle \to \rho} \left\{ \langle \Psi_\rho | \, \mathbf{T}_e + \mathbf{V}_{ee} \, | \Psi_\rho \rangle \right\} + \int v_{ext}(\vec{r}) \rho(\vec{r}) d\vec{r} \right) \tag{1.71}$$

We can then define an universal energy functional that needs to be minimized to yield the exact ground state energy:

$$E[\rho] = \mathcal{F}[\rho] + \int v_{ext}(\vec{r}) \rho(\vec{r}) d\vec{r} \tag{1.72}$$

$$\mathcal{F}[\rho] = \min_{|\Psi_\rho\rangle \to \rho} \left\{ \langle \Psi_\rho | \, \mathbf{T}_e + \mathbf{V}_{ee} \, | \Psi_\rho \rangle \right\} \tag{1.73}$$

and, since its components are related to the kinetic and electronic repulsion operators, we can rearrange Eq. 1.72 further to yield

$$E[\rho(\vec{r})] = T[\rho(\vec{r})] + V_{ee}[\rho(\vec{r})] + \int v_{ext}(\vec{r}) \rho(\vec{r}) d\vec{r} \tag{1.74}$$

$$E[\rho(\vec{r})] = T[\rho(\vec{r})] + J[\rho(\vec{r})] + Q[\rho(\vec{r})] + \int v_{ext}(\vec{r}) \rho(\vec{r}) d\vec{r} \tag{1.75}$$

where the $V_{ee}[\rho]$ term has been further split into a classical Coulomb contribution ($J[\rho]$) and a non-classical part $Q[\rho]$:

$$J[\rho] = \frac{1}{2} \int \frac{\rho(\vec{r}) \rho(\vec{r'})}{|\vec{r} - \vec{r'}|} \tag{1.76}$$

$$Q[\rho] = V_{ee}[\rho(\vec{r})] - J[\rho(\vec{r})] \tag{1.77}$$

Eq. 1.75 indicates how far we can go without introducing extra elements (namely single-particle orbitals) into the DFT framework, so we are left with two unknown terms: the kinetic energy functional $T[\rho]$ and the non-classical term $Q[\rho]$. $T[\rho]$ is larger in magnitude, so early attempts within this orbital-free DFT formalism focused on finding good approximations for this term, largely unsuccessful, as illustrated by the Thomas-Fermi-Dirac (TFD) model. Even later attempts to correct TFD's $T[\rho]$ expression, by including a dependence not only on the density itself, but also on its gradient, hoping to generalize the results for non-uniform systems, did not yield accurate results [57].

1.1.5.2 Orbitals to the Rescue: the Kohn-Sham Framework for DFT

In 1965, Kohn and Sham (KS) proposed introducing the idea of single-particle orbitals to DFT [58]. An intuitive justification is that, in HF theory, the kinetic energy piece of the Hamiltonian seems to be well approximated by the expectation value of a wavefunction composed of products of orbitals. Well enough such that HF can qualitatively, albeit not quantitatively due to lack of correlation effects, describe chemical bonding. Hence, KS-DFT is based on a scheme where a fictitious non-interacting system represented by a Slater determinant that reproduces the exact density of its interacting counterpart is used as an intermediate to approximate some of the quantities in Eq. 1.75. We thus have:

$$\rho(\vec{r}) = \rho_{KS}(\vec{r}) = \sum_i |\phi_i^{KS}(\vec{r})|^2 \tag{1.78}$$

$$\begin{aligned} E[\rho(\vec{r})] = T_s[\rho_{KS}(\vec{r})] + J[\rho_{KS}(\vec{r})] + Q[\rho_{KS}(\vec{r})] \\ + (T[\rho_{KS}(\vec{r})] - T_s[\rho_{KS}(\vec{r})]) + \int v_{ext}(\vec{r})\rho_{KS}(\vec{r})d\vec{r} \end{aligned} \tag{1.79}$$

$$T_s[\rho_{KS}(\vec{r})] = \sum_i \left\langle \phi_i^{KS} \left| -\frac{1}{2}\nabla^2 \right| \phi_i^{KS} \right\rangle \tag{1.80}$$

where $\{|\phi_i^{KS}\rangle\}$, ρ_{KS}, and T_s are, respectively, the fictitious KS single-particle orbitals, density and kinetic energy, the latter being an implicit functional of the density through the KS orbitals. Note that we still do not know how to compute all the terms in Eq. 1.79. We then define the exchange-correlation functional as follows:

$$E_{xc}[\rho(\vec{r})] = Q[\rho(\vec{r})] + (T[\rho(\vec{r})] - T_s[\rho(\vec{r})]) \tag{1.81}$$

and, inspired by the definition of exchange and correlation energy contributions in wavefunction theories, we go a step further and define:

$$E_{xc}[\rho(\vec{r})] = E_x[\rho(\vec{r})] + E_c[\rho(\vec{r})] + T[\rho(\vec{r})] - T_s[\rho(\vec{r})] \tag{1.82}$$

$$E[\rho(\vec{r})] = T_s[\rho(\vec{r})] + J[\rho(\vec{r})] + E_{xc}[\rho(\vec{r})] + \int v_{ext}(\vec{r})\rho(\vec{r})d\vec{r} \tag{1.83}$$

Up to this point, KS-DFT has been discussed as an exact theory. As long as the density of a system can be reasonably represented by a single Slater determinant, such as in the case of single reference chemical systems, we can employ the machinery of KS-DFT. After all of these considerations, it begs the question: can we finally use this arguably powerful tool, capable of compressing the information contained in high dimensional wavefunctions into a simple quantity as the charge density, to calculate energies and properties of complex systems?

Unfortunately, no. At least not without invoking further approximations. We do not, and likely will never, know the exact form of the exchange-correlation functional $E_{xc}[\rho]$. Instead of seeing this as complete failure of KS-DFT, an optimistic aspiring quantum chemist would see this as "glass half full" situation. One that would require even more careful thinking to design physically motivated approximations in order to achieve the elusive goal of accurately computing energies and properties of complex many-particle systems.

1.1.5.3 KS-DFT Alphabet's Soup: Surveying Modern Density Functional Approximations

This is where KS-DFT fledges from its nest of exactness to venture into the wild and vast quantum chemical land of approximations. Over the past few decades, this has been an area of growing interest and fast development. Today, more than a hundred different approximations for E_{xc} have been proposed. This is a testament of how popular KS-DFT is within the quantum chemistry community [59–61]. Naturally, some approximations are better and more accurate than others. Furthermore, it is worth noticing that, contrary to most wave-function theories, it is not guaranteed that these density functional approximations (DFAs) are systematically improvable.

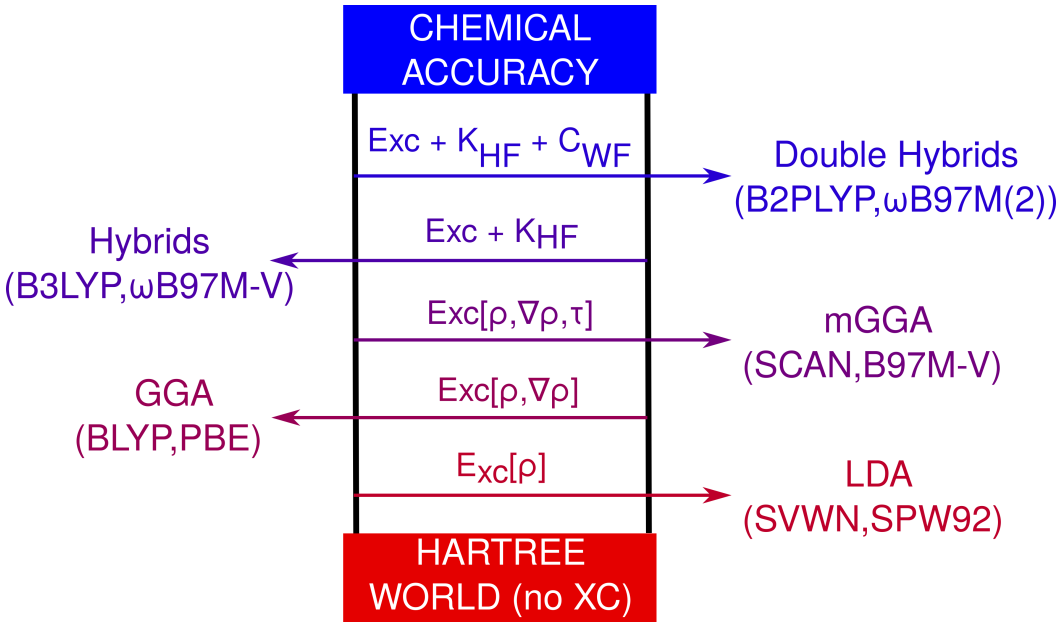


Figure 1.2: Perdew's "Jacob's Ladder" for DFAs, with increasing complexity for the exchange-correlation functional.

John Perdew has introduced the methaphorical idea of a "Jacob's ladder" to group different DFAs within certain categories that contain the same physical ingredients, as shown in Fig. 1.2 [62]. Walking up each rung of the ladder statistically improves that class of DFAs.

The lowest rung is known as the Hartree world, the primeval land devoid of electronic correlation, whereas the highest one is the holy grail of quantum chemistry: the realm of chemical accuracy. In between, the ladder is commonly divided into five rungs, according to the ingredients used in approximating E_{xc}. We do not aim to present an extensive overview of the performance of DFAs in each rung. Instead, we will favor a simplified discussion of the physical content added in each rung and refer the interested reader to check the appropriate references for some of the DFAs mentioned. Finally, we will present expressions for E_{xc} that depend only on the charge density $\rho(\vec{r})$, even though generalizations to spin dependent cases, either through individual $\rho_\alpha(\vec{r})$ and $\rho_\beta(\vec{r})$ for the different spin channels, or one of the components of the magnetization density $m_s(\vec{r}) = \rho_\alpha(\vec{r}) - \rho_\beta(\vec{r})$, can be easily devised. It is worth noting that such extensions are the *de facto* standard in real calculations.

1. **Local Density Approximation (LDA)**: the simplest form of DFA is the one that uses the information about the density at a given point in space:

$$E_{xc} = \int \rho(\vec{r})\epsilon_{xc}(\rho)\ d\vec{r} \tag{1.84}$$

   where ϵ_{xc} is the exchange-correlation energy density. Analytical forms of the exchange functional in LDA can be derived from the uniform electron gas (UEG) [55], while several parametrizations of the correlation energy for UEG have been proposed over the years [63, 64]. The most widely used LDA functionals are the SVWN [65] and SPW92 models [66]. Despite being exact for systems with homogeneous densities, LDA functionals often perform badly in estimating relative energies for a number of processes of chemical interest. For instance, SPW92 has a root-mean square deviation (RMSE) of > 40 kcal/mol for thermochemistry [60] (even though it is undoubtedly an improvement over the "disastrous" TFD model, and even HF theory).

2. **Generalized Gradient Approximation (GGA)**: The next ingredient that can be added to E_{xc} is information about the behavior of the density in the vicinity of a certain point in space. This is accomplished by including the density gradient ($\nabla\rho$) as a dependent variable of ϵ_{xc}

$$E_{xc} = \int \rho(\vec{r})\epsilon_{xc}(\rho, \nabla\rho)\ d\vec{r} \tag{1.85}$$

   GGA functionals are often designed by a careful analysis of exact constraint conditions that they need to satisfy or by fitting against empirical data. The PBE functional [67] is the mostly widely known example of the former category, whereas the BLYP [68] and B97-D3 [69] are examples of empirically fitted functionals. GGAs are often more accurate than LDAs, with PBE's RMSD for thermochemistry being on the order of 10 kcal/mol [60].

3. **Meta-GGA (mGGA)**: The next stop on the highway towards chemical accuracy involves adding even more information about the immediate neighborhood of a point

in space through the density Laplacian $\nabla^2\rho$, or more commonly the kinetic energy density [70], τ, defined as:

$$\tau(\vec{r}) = \sum_i \left|\vec{\nabla}\phi_i(\vec{r})\right|^2 \tag{1.86}$$

Hence, the exchange-correlation piece for a meta-GGA functional can be written as:

$$E_{xc} = \int \rho(\vec{r})\epsilon_{xc}(\rho, \nabla\rho, \tau)\ d\vec{r} \tag{1.87}$$

Popular choices for DFAs at the meta-GGA level include TPSS [71], SCAN [72], both designed to satisfy certain (or all) exact conditions for this class of approximation, and the combinatorially designed B97M-V [73]. mGGAs are usually more accurate than GGAs, with TPSS having a RMSD of about 6 kcal/mol for thermochemistry [60].

4. **Hybrid**: The previous three rungs of the Jacob's ladder use only local (or semi-local) information about the density in order to calculate the exchange-correlation contribution to the total energy of the system. Despite seeing statistical improvement going up each ladder, both LDAs, GGAs and mGGAs suffer from a systematic issue, namely the inexact cancellation of the classical Coulomb contribution ($J[\rho]$) by the local piece of the exchange energy. This leads to the notorious delocalization, or self-interaction (SIE), error in local DFAs [74]. In HF, as previously discussed, we observe exact cancellation of the Coulomb and Exchange, which is a non-local operator, terms for a given electron. Hence, hybrid functionals are designed to account for the mixing of a fraction of non-local HF, or exact exchange, in the exchange-correlation energy expression:

$$E_{xc} = cE_{HF}^K + (1-c)E_x[\rho] + E_c[\rho] \tag{1.88}$$

SIE can be mitigated even further if, instead of adding a fraction of exact exchange globally as in Eq. 1.88, we partitioned the electron-electron repulsion operator into a short-range and a long-range component [75, 76]:

$$\frac{1}{r_{12}} = \frac{1-\text{erf}(\omega r_{12})}{r_{12}} + \frac{\text{erf}(\omega r_{12})}{r_{12}} \tag{1.89}$$

where r_{12} is the interelectronic distance, $\text{erf}(x)$ is the error function, and ω is a parameter that controls the transition between the short (first term on the right-hand side of Eq. 1.89) and long-range (second term on the right-hand side of Eq. 1.89) regimes. Only small fractions of HF exchange are usually added at short-range, as the local piece is, if well designed, capable of describing effects at small interelectronic separation. At long-range, however, HF exchange dominates, usually going to 100% at infinite separation. Popular choices for hybrid functionals include B3LYP [77] and PBE0 [78] as

global hybrids, as well as a mGGA range-separated hybrid, ωB97M-V [79], that has RMSD of about 2.5 kcal/mol for thermochemistry [60]. It is worth noting that other schemes to include exact exchange contributions to the energy functional have been devised, such as local hybrids [80, 81]. Although these have historically received less attention due to computational considerations and certain ambiguity in defining real-space local mixing functions, recent advances have led to the design of computationally efficient and accurate local hybrid approximations [82].

5. **Double Hybrid**: Hybrid functionals add non-local contributions to the exchange piece of E_{xc}. Double hybrids, as the name suggests, add wavefunction contributions to the correlation term [83]. Usually this is accomplished by mixing in a fraction of MP2 energy, as expressed by:

$$E_{xc} = c_1 E_{HF}^K + (1 - c_1)E_x[\rho] + c_2 E^{(2)} + (1 - c_2)E_c[\rho] \tag{1.90}$$

where $E^{(2)}$ is the MP2 correlation energy, and c_1, c_2 are parameters that control mixing between local and nonlocal elements. Being at the top of the Jacob's ladder, double hybrids are expected to be the most accurate DFAs, although at the usually demanding cost of having to evaluate wavefunction based correlation energies. Indeed, a recent double-hybrid, ωB97M(2) presents a remarkable RMSD of 0.96 kcal/mol for thermochemistry [84].

1.1.6 What about excited states?

So far, we have described the tools available within non-relativistic quantum chemistry to calculate the properties and energy of a system in its ground-state. It is worth noticing that, with recent experimental breakthroughs in new spectroscopic techniques [85, 86], and increased interest in harnessing the power of widely available solar energy through materials such as perovskites [87], there is an urgency in developing models that allows us to understand and manipulate electronic excited states of a system.

From a theoretical perspective, we should emphasize that most of the methods previously mentioned can be extended to calculate excited states. For instance, CI theory, relying on exact diagonalization, can yield energies and wavefunctions for excited states, although its truncated version is still plagued by deficiencies in the linear ansatz, leading to an overcorrelation of the ground-state compared to excited states. MP and CC also have their excited state analogs. The algebraic diagrammatic construction (ADC) family of models relies on a polarization propagator that employs the same partitioning scheme as MP theory [88]. On the other hand, equation-of-motion coupled cluster (EOM-CC) [89, 90] theory relies on diagonalization of the similarity transformed electronic Hamiltonian ($\bar{H}$, Eq. 1.65) to yield excitation energies. Even within the DFT framework, time-dependent density functional theory (TDDFT) [91, 92] provides an elegant and incredibly useful avenue to efficiently model most types of excited states.

Our main goal in this section is not to provide an extensive overview of the literature about quantum chemical methods for excited states. It is to highlight the two main approaches that have been commonly used to calculate them. CIS, TDDFT, ADC and EOM-CC theories rely, to some extent, on a truncation (usually at linear order) of a propagator or response function under the effect of a time-dependent perturbation. While useful, these linear-response methods have their limitations, such as not properly describing double or charge transfer excitations in CIS/TDDFT. Although both EOM-CC and TDDFT can be made exact, for instance through expanding the subspace of the diagonalization in the case of the former, or incorporating memory effects in the latter, none of these extensions are, to date, of practical value, as their computational cost would be too demanding to model complex systems.

An alternative approach is to "bend" the variational principle to our needs and "hack" all of the machinery developed so far to land on an (non-Aufbau) excited electronic configuration. This can be accomplished through the use of specialized solvers that will be presented more thoroughly in the later chapters of this work. We will then discuss the use of orbital optimized methods, especially in the context of DFT, to investigate a class of excited states that is poorly described by linear-response models. We highlight that this orbital-optimized, or state-specific, approach is not limited to mean-field theories such as HF and DFT, with extensions to MP [93] and CC [94, 95] already reported in the literature.

1.2 Not so fast! A Brief Overview of Relativistic Quantum Mechanics

The methods and approximations outlined so far are useful to understand a variety of physical and chemical processes. From allowing us to design better pharmaceutical drugs [96], to enabling high-throughput materials discovery [97, 98], non-relativistic quantum chemistry has excelled in modeling, even if qualitatively at times, nature around us. Yet, a few aspects are still missing. For instance, we can not, with the tools so far discussed, understand why gold is a yellow shiny metal, instead of silver-ish/grey like most other transition metals [99, 100]. They can neither explain the physical properties, such as melting and boiling points, of certain coinage metals [101]. In fact, not even spin is fully understood within the framework presented so far, being firstly introduced in an *ad hoc* fashion by Pauli in 1927 to account for the behavior of electrons under the influence of external magnetic fields [102, 103] .

To explain and model these and other phenomena that will be later explored in this work, we need to go beyond the traditional non-relativistic Hamiltonian and find a way to incorporate the effects of special relativity, and the consequences of having a limit for the speed of light into our calculations. The discussion presented in this section is by no means extensive or complete. Instead, we opt to offer a glimpse of how relativity affects the energy levels of simple systems and describe a strategy to generalize some of these findings to more complex, molecular cases. We refer the interested reader to Refs. 104, 105, 106 and 107 for

a more comprehensive discussion.

1.2.1 Particles and Antiparticles: an introduction to the Dirac Equation

Although early attempts by Klein [108] and Gordon [109], reconciling special relativity and quantum mechanics remained an elusive goal. This was due especially to issues in interpreting "negative probabilities" that arise when solving their proposed relativistic equation, which was second order in time. It is important to recall that the non-relativistic time-dependent Schrödinger equation (Eq. 1.1) has a first-order derivative with respect to time, and a second-order spatial derivative through the kinetic energy operator. In 1928, Dirac proposed a new formulation of the relativistic quantum theory [110]. Relativity treats space and time at equal footing, so Dirac, assuming that Eq. 1.1 was valid for a (relativistic) Hamiltonian yet to be revealed, proposed that such Hamiltonian must also have a first order derivative in space.

Dirac's proposed Hamiltonian is of the following form:

$$\mathbf{H} = c\vec{\alpha} \cdot \vec{\mathbf{p}} + c^2\beta \tag{1.91}$$

where $\vec{\mathbf{p}}$ and c are, respectively, the regular momentum operator and the speed of light in vacuum. $\vec{\alpha}$ and β are coefficients later determined by Dirac. Dirac reasoned that the components of $\vec{\alpha}$ and β could not be simple scalars, as that would lead to a preferential spatial orientation. Instead, by requiring that the classical relativistic energy-momentum relation:

$$E^2 = m^2c^4 + c^2p^2 \tag{1.92}$$

is satisfied, we can find a set of equations that $\vec{\alpha}$ and β are required to obey. Namely:

$$\left\{\alpha_k, \alpha_l\right\} = 2\delta_{kl} \tag{1.93}$$

$$\left\{\alpha_k, \beta\right\} = 0 \tag{1.94}$$

$$\left\{\beta, \beta\right\} = 2 \tag{1.95}$$

where $\{A, B\} = AB + BA$ is the usual anticommutator. Eqs. 1.93, 1.95 and 1.95 constitute the Dirac algebra. A careful analysis of these relationships reveals that the simplest case that satisfies the Dirac algebra is if we consider a set of 4×4 matrices. While many representations are possible, a solution to the Dirac algebra is given by the Dirac-Pauli representation:

$$\vec{\alpha} = \begin{bmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{bmatrix} \tag{1.96}$$

$$\beta = \begin{bmatrix} \mathcal{I}_2 & 0 \\ 0 & -\mathcal{I}_2 \end{bmatrix} \tag{1.97}$$

where $\mathcal{I}_2$ is the 2×2 identity matrix, and $\vec{\sigma}$ is the vector of Pauli matrices:

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{1.98}$$

$$\sigma_x = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \tag{1.99}$$

$$\sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{1.100}$$

So far, Eq. 1.91 describes the dynamics of a free relativistic particle represented by a four-component spinor $|\Psi\rangle$. Interpreting the physical content of $|\Psi\rangle$ leads to the notion of any particle (*e.g.*, the electron) having its antimatter, or antiparticle, counterpart (*e.g.*, the positron). For the Dirac spinor $|\Psi\rangle$, two of its components describe the electronic behavior, while the remaining ones concern the dynamics of the positron. Through minimal coupling, that is:

$$\mathbf{H} \rightarrow H + \Phi \tag{1.101}$$

$$\vec{\mathbf{p}} \rightarrow p + \frac{1}{c}\vec{\mathbf{A}} \tag{1.102}$$

where Φ and $\vec{A}$ are the auxiliary scalar and vector potentials, respectively, for the external fields, we are now in position to write down the Dirac Hamiltonian for a particle interacting with its environment (through the auxiliary potentials):

$$\mathbf{H} = c\vec{\alpha} \cdot \left(\vec{\mathbf{p}} + \frac{1}{c}\vec{\mathbf{A}} \right) - \mathbf{\Phi} + c^2\beta \tag{1.103}$$

1.2.2 The Fine Structure Hamiltonian for Hydrogen-like Systems

Undeniably, Eq. 1.103 is remarkable: it describes the formulation of the quantum theory of a particle consistent with special relativity. On the other hand, such a formulation seems to have enlarged the Hilbert space, given that we now have to deal with four-component spinors Ψ. Once again, we should ask ourselves: is all this extra information really necessary to describe the chemical process we are interested? If we constrain ourselves to the realm of low energy physics and chemistry, one in which the velocity of the particle is small fraction of the speed of light in vacuum, then the answer is likely no.

We do not need all of the information contained in the Dirac Equation and its four-component spinor, only pieces of it that are on the relevant energy scale. We can then attempt to expand the Dirac Hamiltonian in powers of $\left(\frac{v}{c}\right)$ to recover the necessary terms and reduce the dimensionality of the Hilbert space. Consequently, this should recover both the fully non-relativistic case and Pauli's *ad hoc* Hamiltonian with a Hilbert space spanned by two-component spinors. This can be accomplished by the so-called Foldy-Wouthuysen (FW) transformation, the details of which can be found elsewhere [104, 111, 112]. It is

worth pointing, nonetheless, that one of the goals of the FW transformation is to decouple the upper and lower two-components pieces of the four-component spinor, allowing us to simply focus in one of them.

For now, we simply note that, after expanding the Dirac Hamiltonian in powers of $\left(\frac{v}{c}\right)$ up to fourth-order and only taking the terms associated with the upper two-components of the Dirac spinor, we obtain the following (for clarity, we reintroduce the factors associated with the mass of the electron and its charge):

$$\mathbf{H} = mc^2 \tag{1.104}$$

$$+ \frac{1}{2m}\left(\vec{\mathbf{p}} + \frac{e}{c}\vec{\mathbf{A}}\right)^2 - e\mathbf{\Phi} \tag{1.105}$$

$$+ g\frac{e}{2mc}\vec{\mathbf{S}} \cdot \tilde{\mathbf{B}} \tag{1.106}$$

$$- \frac{\mathbf{p^4}}{8m^3c^2} + \frac{e}{8m^2c^2}\vec{\nabla} \cdot \vec{\mathbf{E}} + \frac{1}{4m^2c^2}\left(\vec{\mathbf{p}} \times \vec{\mathbf{E}} - \vec{\mathbf{E}} \times \vec{\mathbf{p}}\right) \tag{1.107}$$

$$\mathbf{H} = mc^2 + H_{nr} + H_P + H_{RKE} + H_D + H_{SOC} \tag{1.108}$$

Eqs. 1.104, 1.105, 1.106 recover, respectively, the rest energy of the particle, the non-relativistic Hamiltonian of an electron in an external electromagnetic field (H_{nr}), and Pauli's *ad hoc* description of the interaction of the intrinsic spin of the electron and an external magnetic field (H_P). The new terms (Eq. 1.107) were previously unknown contributions that arise solely from relativistic considerations. In order, the terms in Eq. 1.107 recover corrections for the relativistic kinetic energy (H_{RKE}), considerations about the non-locality of the interaction between the electron and the external electric field through the Darwin term (H_D), and spin-orbit coupling effects (H_{SOC}). For hydrogen-like systems (and more general spherically symmetric potentials), the spin-orbit component reduces to the well-known expression $H_{SOC} \propto \vec{\mathbf{L}} \cdot \vec{\mathbf{S}}$.

Finally, we should note that H_{RKE}, H_D and H_{SOC} lift (some of) the degeneracies of the non-relativistic energies of hydrogenic systems (Eq. 1.14). Moreover, H_{SOC} does not commute with either $\vec{\mathbf{L}}$ or $\vec{\mathbf{S}}$ operators, and hence coupling between states with different angular momentum and spin orientations is possible. This requires the definition of a new quantum number based on the coupled/total angular momentum operator $\vec{\mathbf{J}} = \vec{\mathbf{L}} + \vec{\mathbf{S}}$ to label the states of the relativistic electronic Hamiltonian. Through perturbation theory, we can obtain the energies of the j-label states of hydrogenic systems:

$$E_{nj} = \left(-\frac{Z^2}{2n^2}\right)\left[1 - \left(\frac{Z}{cn}\right)^2\left(\frac{3}{4} - \frac{n}{j+\frac{1}{2}}\right)\right] \tag{1.109}$$

Curiously, we can see that Eq. 1.109 still predicts degenerate energies for the $2s_{\frac{1}{2}}$ and $2p_{\frac{1}{2}}$ of hydrogen-like systems. Lifting this degeneracy requires considering the role of vacuum fluctuations of the photon field within the context of the Lamb shift [113, 114]. Despite being a remarkably rich and interesting field that has gathered increasing interest from the quantum chemistry community in recent years, quantum electrodynamics [115–118] and its extension to molecular systems [119] is out of the scope of this work.

1.2.3 Quantum Chemistry and Relativity

Up to now, our discussion about relativistic effects has been limited to a single particle (or the electron-positron pair described by the four-component Dirac spinor). To treat many-particle molecular systems, we need to find a way to extend the Dirac Hamiltonian. As was done in the non-relativistic case, we can decompose the full relativistic Hamiltonian into its one and two-body components. Intuitively, the one-body relativistic Hamiltonian for the molecular case shall be written as a sum of single-particle Dirac Hamiltonians (Eq. 1.103). The two-body Hamiltonian, however, requires more careful consideration. What two-body interactions should be included?

Naturally, we should include the Coulombic V_{ee} term (which we will now denote as $V_{ee}^{C} = \frac{1}{r_{ij}}$). However, this assumes that electrons interact instantaneously. This means that if one electron at one side of the universe moves, another one an the other side of the cosmos will react immediately. While exaggerated, the previous statement illustrates that the Coulomb repulsion term breaks special relativity. In order to correct this, we need to include the so-called Breit interaction in the Hamiltonian [120] (here just written for the interaction between two electrons for simplicity):

$$V_{ee}^{B} = -\frac{1}{2r_{ij}}\left[\vec{\alpha}_i \cdot \vec{\alpha}_j + \frac{(\vec{\alpha}_i \cdot \vec{r}_{ij})\,(\vec{\alpha}_j \cdot \vec{r}_{ij})}{r_{ij}^2}\right] \tag{1.110}$$

where $\vec{\alpha}_i$ is the vector of Dirac-Pauli matrices (Eq. 1.96) for particle i. Eq. 1.110 include both the interactions between the magnetic moments of two-electrons, as well as retardation effects. Hence, the fully relativistic molecular Hamiltonian can be written as:

$$\mathbf{H} = \sum_i \mathbf{h}_i + \frac{1}{2}\sum_{i,j} \mathbf{V}_{ij}^{C} + \mathbf{V}_{ij}^{B} \tag{1.111}$$

where $\mathbf{h}_i$ and $\mathbf{V}_{ij}^{B}$ are given by Eq. 1.103 and Eq. 1.110, respectively.

With the Hamiltonian at hand, quantum chemists have tried to extend most of the electronic structure machinery previously discussed to the relativistic realm [121]. Once again, we should note that the Hamiltonian in Eq. 1.111 acts on the space of four-component (many-particle) wavefunctions, and therefore contains more information than is necessary for typical problems of chemical interest. Several approximations, inspired by the FW transformations previously discussed, have been devised in an attempt to decouple the electronic and positronic blocks of the relativistic Hamiltonian. Among them, we highlight the Breit-Pauli (BP) Hamiltonian [122], the Douglas-Hess-Kroll (DKH) [123–126] and the exact-two component method (X2C) [127, 128].

Finally, we shall briefly discuss some of the details of X2C. The starting point for deriving the X2C relativistic model is the four-component one-electron Dirac Hamiltonian (Eq. 1.112) represented in a restricted kinetic balance (RKB) form [129]. Our goal is to use a unitary transformation that effectively decouples the positive and negative-energy solutions of the Dirac equation, since we are only interested in describing electrons. In Eq. 1.112, T, V

and S are the usual non-relativistic kinetic energy, nuclear attraction and overlap matrices, respectively, represented in a basis of atomic orbitals $\{\phi_\mu\}$. W is defined as the matrix representation of the operator in Eq. 1.113, where $\vec{\sigma}$ is the vector of Pauli matrices, and $\vec{p}$ and V are the momentum and nuclear-attraction operators, respectively. As indicated in Eq 1.113, $\hat{W}$ can be decomposed into spin-free ($\hat{W}_{\mathrm{SF}}$) and spin-orbit ($\hat{W}_{\mathrm{SO}}$) components (likewise, the matrix representation W can also be separated into W_{SF} and W_{SO}. Finally, the solutions of Eq. 1.112 are characterized by their large (C_L) and small (C_S) components.

$$\begin{bmatrix} V & T \\ T & \frac{W}{4c^2} - T \end{bmatrix} \begin{bmatrix} C_L \\ C_S \end{bmatrix} = E \begin{bmatrix} S & 0 \\ 0 & \frac{1}{2c^2}T \end{bmatrix} \begin{bmatrix} C_L \\ C_S \end{bmatrix} \tag{1.112}$$

$$\begin{aligned} \hat{W} &= (\vec{\sigma}\cdot\vec{p})V(\vec{\sigma}\cdot\vec{p}) \\ &= (\vec{p}\cdot V\vec{p} + i\vec{\sigma}\cdot(\vec{p}\times V\vec{p})) \\ &= \hat{W}_{\mathrm{SF}} + i\vec{\sigma}\cdot\hat{W}_{\mathrm{SO}} \end{aligned} \tag{1.113}$$

$$W_{\mu\nu} = \langle\phi_\mu|\,\hat{W}\,|\phi_\nu\rangle \tag{1.114}$$

We can obtain the coupling matrix X for the positive energy solutions of Eq. 1.112 as the ratio between the large and small components, as indicated in Eq. 1.115. The renormalization matrix R is then defined as in Eq. 1.116, where $\tilde{S}$ (Eq. 1.117) is a modified overlap matrix that takes into account the folding of the small component into the large one.

$$X = C_S(C_L)^{-1} \tag{1.115}$$

$$R = S^{-1/2}\left(S^{-1/2}\tilde{S}S^{-1/2}\right)^{-1/2}S^{1/2} \tag{1.116}$$

$$\tilde{S} = S + X^\dagger\frac{1}{2c^2}X \tag{1.117}$$

With the X and R matrices, we can now calculate the effective X2C kinetic energy (Eq. 1.118) and nuclear attraction (Eq. 1.119) operators for subsequent electronic structure calculations. Slight modifications are still necessary to incorporate these effects into the two-body terms[130, 131], but this is out of the scope of the present work. As Eq. 1.113 indicates, a separation of scalar and vector (spin-orbit) effects can be naturally accomplished within X2C.

$$T_{X2C} = R^\dagger\left(TX + X^\dagger T - X^\dagger TX\right)R \tag{1.118}$$

$$V_{X2C} = R^\dagger\left(V + \frac{1}{4c^2}X^\dagger WX\right)R \tag{1.119}$$

1.3 The Three Pillars of Modern Quantum Chemistry

It has been a long theoretical journey so far. From discussing the basis of Schrödinger's exact quantum mechanics, to the main approximations used in non-relativistic quantum chemistry, and even beyond the borders of conventional electronic structure theory, reaching a "shadowy" land where the speed of light is finite. This is deliberately an incomplete, and, at times, biased discussion . Our goal was to present the reader with an overview of the main tools necessary to, hopefully, interpret the next chapters of this work. The primary message that we wished to convey is that quantum mechanics is one of the great monuments of modern science. Applying it to complex systems is, as prefaced by Dirac back in 1929, a Herculean task. Physically motivated approximations, as well as better and more efficient algorithms are necessary to advance its predictive power.

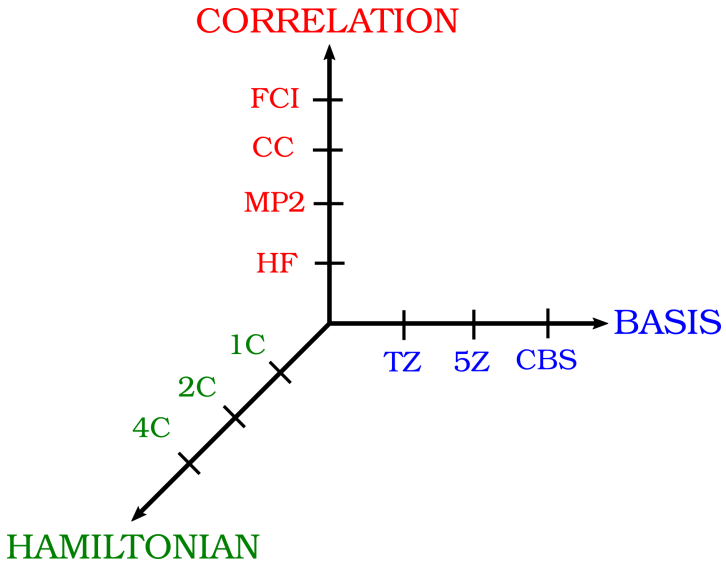


Figure 1.3: The three pillars of modern quantum chemistry: advancing basis set design, development of new approximations to recover electronic correlation and augmenting the number of components of the Hamiltonian to describe relativistic effects.

Thus, it is instructive to illustrate the three main thrusts upon which one can act to accomplish these goals (Fig. 1.3). First, we acknowledge that, while omitted from the previous discussion, electronic structure calculations are usually carried out in a finite basis. Plane waves [132], real-space grids [133], and more commonly single-particle Gaussian functions (or atomic orbitals) [134] are usual choices. Designing more efficient and compact basis sets are of paramount importance to reduce computational requirements (such as memory), although this topic has not been explored in the present work.

Second, we need to understand what are the limitations of current correlated methods, either based on wavefunction or density functional theories, to be able to recover as much

correlation effects as possible. Modern approaches attempt to accomplish this by choosing appropriate orbitals [135], regularizing certain quantities hoping to incorporate higher order correlation effects [136], or levering the power of machine learning to design better density functional approximations [137].

Lastly, we consider the "dimensions" of the Hamiltonian we are interested in. In the non-relativistic case (or one that simply incorporates scalar effects), a single-component wavefunction is usually enough. By single-component we mean that there is no coupling between different spin channels. On the other hand, a fully relativistic treatment requires a four-component spinor, which contains a lot more information than necessary for most chemical applications. A compromise seems to be the development of new approximations within the two-component framework, such as X2C, where some heavy lifting is done *a priori* to decouple the components of interest, usually the ones associated with electrons, and then proceed with the application of the usual methods of non-relativistic quantum chemistry to calculate the energies and properties of complex chemical systems.

1.4 Outline

The remainder chapters of this thesis aim to advance certain aspects regarding the last two of the three main pillars of modern quantum described in the previous section. A brief synopsis of each chapter follows.

Chapter 2

The content and figures of this chapter are reprinted or adapted with permission from **Cunha, L. A.**; Lee, J.; Hait, D.; McCurdy, C. W.; Head-Gordon, M. "Exploring spin symmetry-breaking effects for static field ionization of atoms: Is there an analog to the Coulson–Fischer point in bond dissociation?" *J. Phys. Chem.*, **155**, 014309, 2021.

Löwdin's symmetry dilemma is a ubiquitous issue in approximate quantum chemistry. In the context of Hartree-Fock (HF) theory, the use of Slater determinants with some imposed constraints to preserve symmetries of the exact problem may lead to physically unreasonable potential energy surfaces. On the other hand, lifting these constraints leads to the so-called broken symmetry solutions that usually provide better energetics, at the cost of losing information about good quantum numbers that describe the state of the system. This behavior has been previously extensively studied in the context of bond dissociation. This paper studies the behavior of different classes of Hartree-Fock spin polarized solutions (restricted, unrestricted, generalized) in the context of ionization by strong static electric fields. We find that, for simple two-electron systems, UHF is able to provide a qualitatively good description of states involved during the ionization process (neutral, singly-ionized and doubly ionized states), whereas RHF fails to describe the singly ionized state. For more complex systems, even though UHF is able to capture some of the expected characteristics of

the ionized states, it is constrained to a single M_s (diabatic) manifold in the energy surface as a function of field intensity. In this case a better qualitative picture can be painted by GHF as it is able to explore different spin manifolds and follow the lowest solution due to lack of collinearity constraints on the spin quantization axis.

Chapter 3

The content and figures of this chapter are reprinted or adapted with permission from Roychoudhury, S.; **Cunha, L. A.**; Head-Gordon, M.; Prendergast, D.; "Changes in polarization dictate necessary approximations for modeling electronic deexcitation intensity: Application to x-ray emission" *Phys. Rev. B*, **106**, 106, 2022.

Accurate simulation of electronic excitations and de-excitations are critical for complementing complex spectroscopic experiments and can provide validation to theoretical approaches. Using a generalized framework, we contrast the accuracy and validity of orbital-constrained and linear-response approaches that build upon Kohn-Sham density functional theory (DFT) to simulate emission spectra of electronic origin and propose a new and efficient approximation, named Many-Body X-ray Emission Spectroscopy or MBXES, for simulating such processes. We show analytically as well as with computed examples that for electronic (de)-excitation leading to an appreciable change in polarization (i.e., density rearrangement), the adiabatic approximation in a response-based formalism will be inadequate for the calculation of oscillator strength. Thus, such a change (e.g. in the net electrostatic dipole moment of a finite system) can be used as a metric for evaluating the applicability of the adiabatic response-based approach and can be particularly valuable in X-ray emission spectroscopy. On the other hand, MBXES, the flexible method introduced in this article, can compute oscillator strengths accurately at a much lower computational expense on the basis of two DFT based self-consistent field calculations. Using illustrative examples of emission spectra, the efficacy of the MBXES method is demonstrated by comparison with its parent theory, orbital-optimized DFT, and with experiments.

Chapter 4

The content and figures of this chapter are reprinted or adapted with permission from **Cunha, L. A.***; Hait, D.*; Kang, R.; Mao, Y.; Head-Gordon, M. "Relativistic Orbital-Optimized Density Functional Theory for Accurate Core-Level Spectroscopy" *J. Phys. Chem. Lett.*, **13**, 15, 3438-3449, 2022. (* indicates equal contribution from authors.)

Core-level spectra of 1s electrons (K-edge) of elements heavier than Ne show significant relativistic effects. We combine recent advances in orbital optimized density functional theory (OO-DFT) with the spin-free exact two-component (X2C) model for scalar relativistic effects, to study K-edge spectra of elements in the third period of the periodic table. OO-DFT/X2C is found to be quite accurate at predicting energies, yielding ~ 0.5 eV root

mean square error (RMSE) vs experiment with the local SCAN functional and the related SCANh hybrid functional. This marks a signficant improvement over the > 50 eV deviations that are typical for the popular time-dependent DFT (TDDFT) approach. Consequently, experimental spectra are quite well reproduced by OO-DFT/X2C, without any need for empirical shifts for alignment between the two. OO-DFT/X2C therefore is a promising route for computing core-level spectra of third period elements, as it combines high accuracy with ground state DFT cost. We also explored K and L edges of 3d transition metals to identify possible limitations of the OO-DFT/X2C approach and discuss what additional features would be needed for accurately modeling the spectra of such electrons.

Chapter 5

The content and figures of this chapter are reprinted or adapted with permission from Carter-Fenk, K.; **Cunha, L. A.**; Arias-Martinez, J. E.; Head-Gordon, M. "Electron-Affinity Time-Dependent Density Functional Theory: Formalism and Applications to Core-Excited States" *J. Phys. Chem. Lett.*, **13**, 41, 9664-9672, 2022.

The particle-hole interaction problem is longstanding within time-dependent density functional theory (TDDFT) and leads to extreme errors in the prediction of K-edge X-ray absorption spectra (XAS). We derive a linear-response formalism that uses optimized orbitals of the n–1-electron system as reference, building orbital relaxation and a proper hole into the initial density. Our approach is an exact generalization of the static-exchange approximation that ameliorates particle-hole interaction error associated with the adiabatic approximation and reduces errors in TDDFT XAS by orders of magnitude. With a statistical performance of just 0.5 eV root-mean-square error and the same computational scaling as TDDFT under the core-valence separation approximation, we anticipate that this approach will be of great utility in XAS calculations of large systems.

Chapter 2

Exploring Spin Symmetry-Breaking Effects for Static Field Ionization of Atoms

2.1 Introduction

Discussions on the usefulness of symmetry-broken approximate solutions are a familiar topic in time-independent quantum chemistry [138–143]. The issue is encapsulated in what Löwdin has called the "symmetry dilemma" [144]: a more flexible trial wavefunction that does not necessarily preserve all of the symmetries of the exact Hamiltonian might lead to better energetics (*i.e.* lower energy on account of the variational principle) at the cost of losing good quantum numbers that characterize the state of a given system. Within the single determinant Hartree-Fock (HF) model, classification of these symmetry-broken solutions is based on group theory considerations [145], but most commonly in electronic structure one uses terminology that reflects constraints imposed on the orbitals the comprise the single determinant [146]. For instance, requiring that both α and β spin-orbitals share a common set of spatial functions (i.e. the electrons are paired whenever possible) leads to the well-known restricted closed-shell ($M_s = 0$) Hartree-Fock (RHF) and the more general ($M_s \neq 0$) restricted open-shell HF (ROHF) models. RHF and ROHF are both eigenstates of total spin, $\hat{S}^2$, and its z component, $\hat{S}_z$. Lifting this spin pairing constraint, such that α or β spin orbitals can have different spatial functions, gives us the unrestricted Hartree-Fock (UHF) model, whose wavefunction is no longer an eigenfunction of $\hat{S}^2$. Further symmetry lowering by abolishing the notion of separate sets of α and β orbitals leads to the generalized Hartree-Fock (GHF) wavefunction, which is not an eigenstate of either $\hat{S}^2$ or $\hat{S}_z$. Furthermore, number symmetry can also be relaxed following Hartree-Fock Bogoliubov theory [147, 148], which yields a state that is not an eigenstate of the particle number operator ($\hat{N}$).

Within a finite basis, the HF energy, $E_{\mathrm{HF}}(\boldsymbol{\theta})$, is a function of orbital rotation parameters, $\boldsymbol{\theta}$ that mix occupied and virtual orbitals [24, 25, 149]. A solution of the Hartree-Fock

equations zeros the orbital rotation gradient, such that $\nabla_{\boldsymbol{\theta}} E_{\mathrm{HF}} = \mathbf{0}$ (i.e. it is guaranteed to be a stationary point). However a solution is not necessarily a minimum, and an analysis of the orbital Hessian, $E^{\boldsymbol{\theta\theta}}_{\mathrm{HF}}$, is required to characterize the nature of a stationary point. If all of the eigenvalues of the Hessian are greater than zero, we have found a solution that is a local minimum, and is said to be stable. The question of which set of orbital parameters to include in the evaluation of the Hessian arises[150]: if the HF solution is stable within the manifold determined by certain symmetry constraints (e.g. the RHF constraints), it is said that such a solution is internally stable. Lifting some or all of these symmetry constraints leads to characterization of the stationary point in manifolds of higher dimension than the one it was originally optimized in (e.g. characterizing an RHF solution in the space of UHF variations). This corresponds to an analysis of the external stability of the given solution. It might be useful to notice that stability analysis is closely connected with TDDFT/TDHF linear response equations [91, 143] which are routinely available in standard quantum chemistry packages.

For common problems in ground-state electronic structure, internal stability analysis avoids convergence to spurious saddle points and excited states. On the other hand, external stability analysis sometimes reveals interesting physical insight into the nature of electron correlation. For instance, when single bonds are stretched beyond the so called Coulson-Fischer (CF) point [151], the lowest triplet excited state (T_1) starts to mix with the singlet ground state (S_0) in a process known as the "triplet instability" [152, 153], which leads to spin symmetry breaking. The resulting spin polarized UHF state has lower energy than the RHF state, but is no longer an eigenstate of $\hat{S}^2$: the state is said to be spin-contaminated due to its mixed singlet-triplet character. The RHF solution that pairs electrons in order to keep a well-defined $\hat{S}^2$ eigenstate fails to provide a qualitatively correct description of the dissociation process. The RHF pairing constraint, which fixes natural occupation numbers at 0 or 2 even at the dissociation limit (when one would expect the two orbitals to be singly occupied) leads to a state that spuriously preserves ionic character, which consequently fails to approach the correct asymptotic limit at complete dissociation. On the other hand, UHF is able to successfully provide a qualitatively accurate description of the ground state potential energy surface for single bond dissociation [24]. Similar considerations apply to molecules that are singlet diradicaloid [154–156] in character: because two orbitals have occupation numbers significantly different from two and zero, RHF cannot be qualitatively correct, while UHF exhibits spin contamination due to mixing of S_0 and T_1.

There are surely other situations in which similar symmetry dilemma arises, but where its consequences have not been so thoroughly explored as bond dissociation processes and singlet diradicaloids. One example lies in the interaction of atoms and molecules with strong electric fields [157–165]. Strong field chemistry and physics is a rapidly developing field, because of the rich array of new highly non-linear phenomena that emerge. For example, a strong oscillating field will nearly ionize bound electrons in one direction in a first half-cycle followed by reattachment and near ionization in the opposite direction in the second half-cycle. This leads to high harmonic generation (HHG) [166–170], the emission of radiation by the driven bound system at frequencies which are many times that of the applied

radiation. HHG and related phenomena are building blocks for the new field of attosecond science [171–174]. Therefore the use of quantum chemistry methods to model strong field phenomena is also attracting increasing interest [175–180]. However, to date there has been no systematic exploration of the role of symmetry-breaking in the mean field HF method in this context, to our knowledge. This work represents a first step in this direction, though the issue of inadequate field-free Hartree-Fock reference for real-time methods in strong field environments has been briefly mentioned in some previous works [181–184] .

By definition strong fields are those whose scale approaches the strength of internal electric fields experienced by the valence electrons of atoms and molecules (e.g. 1 a.u. for the H atom). Therefore the molecule and the field must be considered as a combined system, rather than treating the field as a perturbation. If such fields are static, then strong field ionization becomes possible, and some electrons may be unbound. Yet because the molecule-field system is treated as a whole, the symmetry dilemma may arise for at least some values of the applied field. In this sense, the magnitude of the field is a control variable similar to the degree of bond-stretching in the dissociation of a closed-shell stable molecule. This paper explores the role of symmetry-breaking in HF solutions for atom-field systems where a static field is scanned across values that can strip one or more electrons from the atom. This investigation is interesting in its own right, and also may help to set the stage for a subsequent analysis in the context of time-dependent fields. This chapter is arranged as follows: in Section 2.2 we discuss our approach to approximate the description of continuum-like states using ghost basis functions, an analysis of the HF spin symmetry-broken solutions for the static field ionization of helium and neon is presented in Sections 2.3 and 2.4, respectively, and in Section 2.5 we assess our main conclusions while presenting an outlook of future work.

2.2 Static Field Ionization in a Finite Basis

The first issue that arises when applying an electric field to an atomic or molecular system is that, strictly speaking, bound states are no longer supported by the combined potential arising from the system and the field [185, 186]. These states are now resonances and a suitable discretization for both localized and continuum states are needed to describe them. From a theoretical perspective, the scattering community has developed a wide range of tools to treat these resonance states, such as exterior complex scaling (ECS) [187–194], complex absorbing potentials (CAP) [195–203], and different grids as schemes to discretize space [204–210]. On the other hand, the quantum chemistry community has long advocated for the use of atomic orbital (AO) expansions based on Gaussian functions as an efficient way to compute ground and excited state properties of molecular systems. Since our goal is to study static field ionization exploring the common toolbox of quantum chemistry, we shall resort to the usual finite AO basis set treatment. What are the consequences of such choice? The most concerning one is that ionization losses are hindered, since electronic density will be constrained near the atom/molecule. Two possible ways [211] of circumventing this issue are the use of highly diffuse basis sets and discretizing the relevant part of the space by

adding a ghost atoms/functions that could be populated by electron density as ionization takes place. We have opted for this second option and added a series of ghost functions in a line passing through the atom and in the direction of the applied field. This allows us to capture some of the effects of ionization as electron density can escape the system and populate the discretized space (Fig. 2.1). This setup is inspired by previous works that aimed to study the real-time electron dynamics of systems in the context of strong fields and high-harmonic generation (HHG), where the addition of diffuse and ghost functions to the basis set is important for a good description of the Rydberg and unbound states of the system, respectively [211].

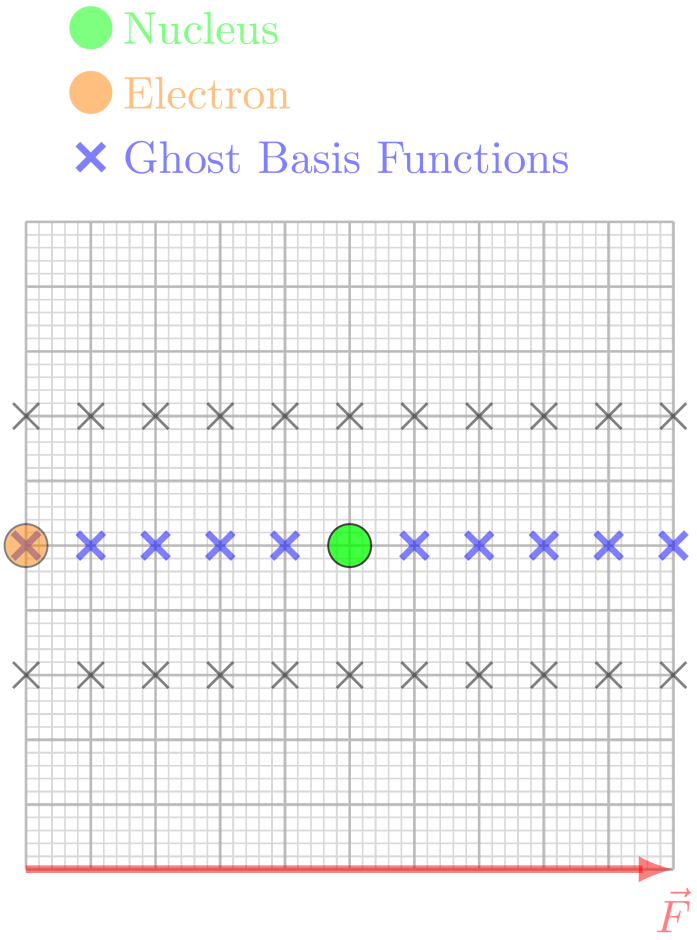


Figure 2.1: Setup of our model for static field ionization. A series of ghost functions was added passing through the atom and in the direction of the field, allowing electron density to escape/ionize from the system (the grey crosses represent additional race track ghost functions, as discussed in Sec. 2.4).

Given this setup for our problem, we also need to interpret what the exact results would

look like in this finite basis set approach. We start our analysis by noticing that, assuming an adequate discretization of the space by the ghost centers, we can separate the problem/potential into two parts: one associated with the molecular/atomic system and another related to the artificial box that is effectively created by the extent of the ghost centers. Introductory quantum mechanics teaches us that, to first order, the energy levels of a charged particle in a box subjected to a uniform static electric field decrease linearly with the strength of the field [212]. Moreover, increasing the field strength affects the molecular potential by suppressing the barrier to ionization for higher excited state (especially Rydberg states) and increasing the tunneling probability for low-lying states. In this sense, higher excited states easily become continuum-like states and "dive" down in energy as the field becomes stronger. At certain field strengths, these continuum-like states start to interact with the low-lying bound states and an avoided crossing is observed: the former bound state acquires continuum-like character by localizing around the edge of the artificial box, with an energy that decreases linearly as the field strength continues to increase. Notice that the applied electric field turns all of the states of the system into resonances with finite lifetime [212], and a proper discussion on how to obtain DC Stark lifetimes for several molecular system using quantum chemistry methods has already been presented elsewhere [177, 178, 213].

Fig. 2.2 illustrates this process for a He atom (described by the functions contained in the aug-cc-pVTZ basis set[214, 215]) in an artificial box comprised of 18 hydrogen STO-3G[216] s-type ghost centers spaced by 0.5 Å spanning a range of 4.5 Å around the atom (which should be adequate for an unambiguous determination of the state of ionization of the system given that this distance is much larger than the atomic radius of He). We shall focus on the behavior of the ground-state S_0 and the first singlet excited state S_1 as a function of field strength. Initially, in the absence of the field, these states are well separated in energy and there is no mixing between them. As the field strength increases, we observe the characteristic quadratic polarization effect on the energy (which can be described by perturbation theory) [217]. Higher lying states are more polarizable, so this effect becomes more evident. These high energy excited states are also more easily ionized and the transition between the quadratic behavior characteristic of a bound state to a linear dependence of the energy on field strength happens at lower fields. At F = 0.175 a.u., we observe an avoided crossing between the first singlet state (which has been previously ionized) and the unionized ground state. This leads to a change in character of the ground state after the avoided crossing: initially unionized, with both electrons localized around the atom, the ground state becomes singly ionized, with one electron detaching from the atom and localizing at the edge of the box. The second ionization is characterized by yet another avoided crossing between the singly ionized S_0 and doubly ionized S_1 at a field of F = 0.36 a.u. Between these two field strengths, we note that the singly ionized T_1 state is degenerate with the ionized ground state S_0. It is worth noticing that the potential energy curves presented here as a function of field strength are very similar to stabilization graphs proposed by Simons [218] and widely used to describe temporary ions and other atomic/molecular resonances [219]. There, information about the resonance's lifetime can be obtained by an analysis of the avoided crossing that arise by changing the variational parameters associated with the

basis set. Here, however, the basis set is fixed and the energy is plotted as a function of varying field strength.

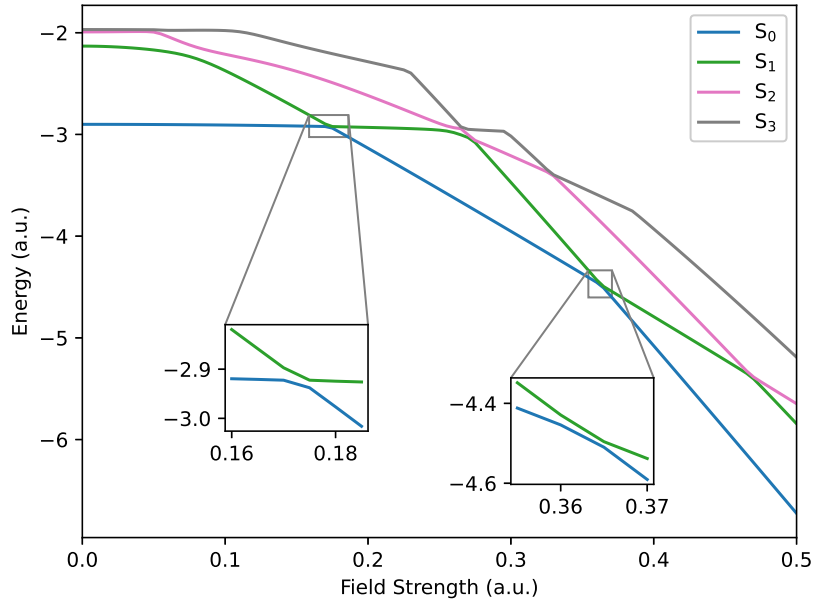
Our goal is to understand how different approximations, such as HF and MP2 can recover this exact finite basis behavior. We shall put emphasis on the effects of allowing for spin polarization (*i.e.*, difference between using restricted and unrestricted orbitals) to describe field ionization as an analog to what happens in bond dissociation of molecules. All calculations were performed with the Q-Chem 5.3 package [220] following the same protocol: we used the aug-cc-pVTZ basis set for the atomic center and hydrogenic STO-3G s functions for the ghost centers. We also performed internal stability analysis in order to ensure that the solution within the domain of each HF class (RHF, UHF and GHF) was the lowest possible.

2.3 Field Ionization of He

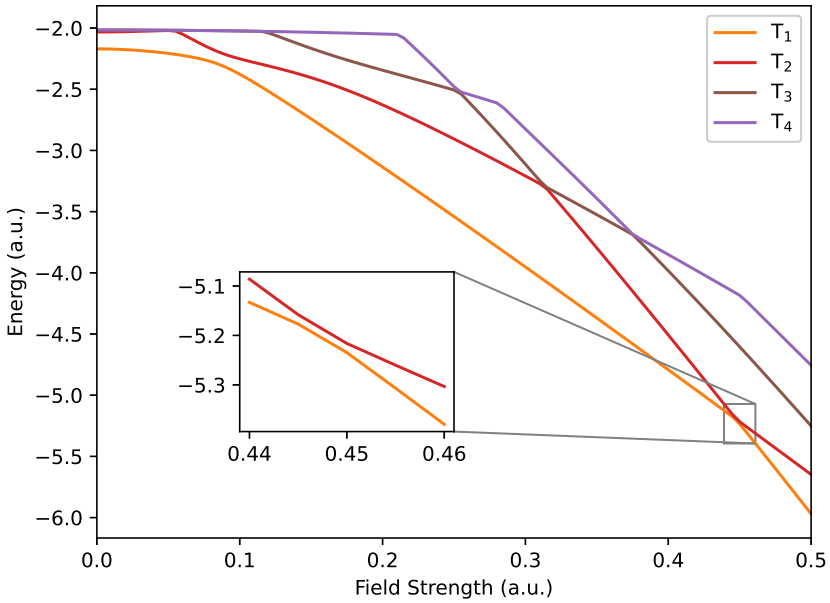
2.3.1 Hartree-Fock (HF) Minimum Basis Model

We start our discussion about the relation between spin polarization and static field ionization by analyzing the behavior of the Hartree-Fock (HF) solution in a simple, yet instructive, minimum basis model of a two electron system. This model is comprised of a single basis function localized around the atom (represented by $|A\rangle$ in Table 2.1) and a single function (represented as $|C\rangle$ in Table 2.1) placed at a distance d from the atom in the direction of the applied field to represent the discretized continuum and that can support the flux of ionized electrons once the field is switched on. It should be noted that this is analogous to the toy model used to describe the bond stretching of H_2 in the minimum basis [24, 143]. Here, we have also assumed that the distance d is large enough such that there is no overlap between the orbitals representing the atom and the discretized continuum. Moreover, we assumed that the ghost function is diffuse enough that we can neglect its kinetic energy ($\langle C|\hat{T}|C\rangle \approx 0$), but not too diffuse relative to the distance d. We can then construct the HF orbitals as linear combinations of these two functions.

In this toy model, the RHF solution is controlled by a single variational parameter θ that mixes the atomic basis function with the continuum-like ghost function (Table 2.1). Analyzing the stationary and stability conditions of the energy obtained as the expectation value of the Hamiltonian with the RHF determinant leads to two limiting cases. The solution characterized by the double occupancy of the atom-centered orbital ($\theta = 0$) is found to be stable when the product of the applied field strength (F) and the distance between the basis functions (d) is less than the IP_1, the first ionization potential of the system (*i.e.* $Fd < IP_1$). On the other hand, the $\theta = \frac{\pi}{2}$ solution, which corresponds to 2 electrons paired up on the ghost function that supports the ionized flux, is stable when $Fd > IP_2$, where IP_2 is the second ionization potential of the system. At intermediate field strengths, $IP_1 < Fd < IP_2$, the stable solution is given by a state that has mixed neutral and doubly ionized character. It is interesting to notice that RHF only allows for the combined ionization of the electron pair and the intermediate state associated with single ionization is never reached. Hence, we



(a) Singlet states



(b) Triplet states

Figure 2.2: FCI (exact) results for the static field ionization of He for (a) lowest four singlet states, (b) lowest four triplet states. The "kinks" correspond to avoided crossings between continuum-like states associated to the artificial finite box and localized states from the atom. For the ground state, the crossings also give us the signature associated to ionization events. The atom is described by the aug-cc-pVTZ basis set, whereas each ghost center is described by hydrogenic STO-3G 1s function separated by 0.5 Å.

see that restricted HF (RHF) cannot recover the exact behavior shown in Fig. 2.2 for the ground state, as it only connects the neutral solution (with the electrons paired up on the atom) and the doubly ionized state.

Model Assumptions		
1a. $\|A\rangle$ - atomic spatial orbital 1b. $\|C\rangle$ - continuum-like orbital 2. $\langle A\|C\rangle \approx 0$ 3. $\langle C\|\hat{T}\|C\rangle \approx 0$		
Spatial Orbitals		
$\|1\rangle = \cos\theta_1 \|A\rangle + \sin\theta_1 \|C\rangle$ $\|2\rangle = \cos\theta_2 \|A\rangle + \sin\theta_2 \|C\rangle$		
Stable RHF - $\|\Psi\rangle = \|1\bar{1}\rangle$		
$\theta = 0$ $Fd < IP_1$ $\|\Psi\rangle = \|A\bar{A}\rangle$	$0 < \theta < \frac{\pi}{2}$ $IP_1 < Fd < IP_2$ mixed solutions	$\theta = \frac{\pi}{2}$ $Fd > IP_2$ $\|\Psi\rangle = \|C\bar{C}\rangle$
Stable UHF - $\|\Psi(\theta_1,\theta_2)\rangle = \|1\bar{2}\rangle$		
$\theta_1 = \theta_2 = 0$ $Fd < IP_1$ $\|\Psi\rangle = \|A\bar{A}\rangle$	$\theta_1 = 0, \theta_2 = \frac{\pi}{2}$ or $\theta_1 = \frac{\pi}{2}, \theta_2 = 0$ $IP_1 < Fd < IP_2$ $\|\Psi\rangle = \|C\bar{A}\rangle$ or $\|\Psi\rangle = \|A\bar{C}\rangle$	$\theta_1 = \theta_2 = \frac{\pi}{2}$ $Fd > IP_2$ $\|\Psi\rangle = \|C\bar{C}\rangle$

Table 2.1: Analytic representation of molecular orbitals and ionization states predicted by HF for minimal basis model. $|A\rangle$ and $|C\rangle$ are the atomic and ghost/continuum basis functions, respectively. The stationary and stability conditions for RHF and UHF lead to different character of the solutions as a function of field strength and UHF is the only model capable of describing single ionization.

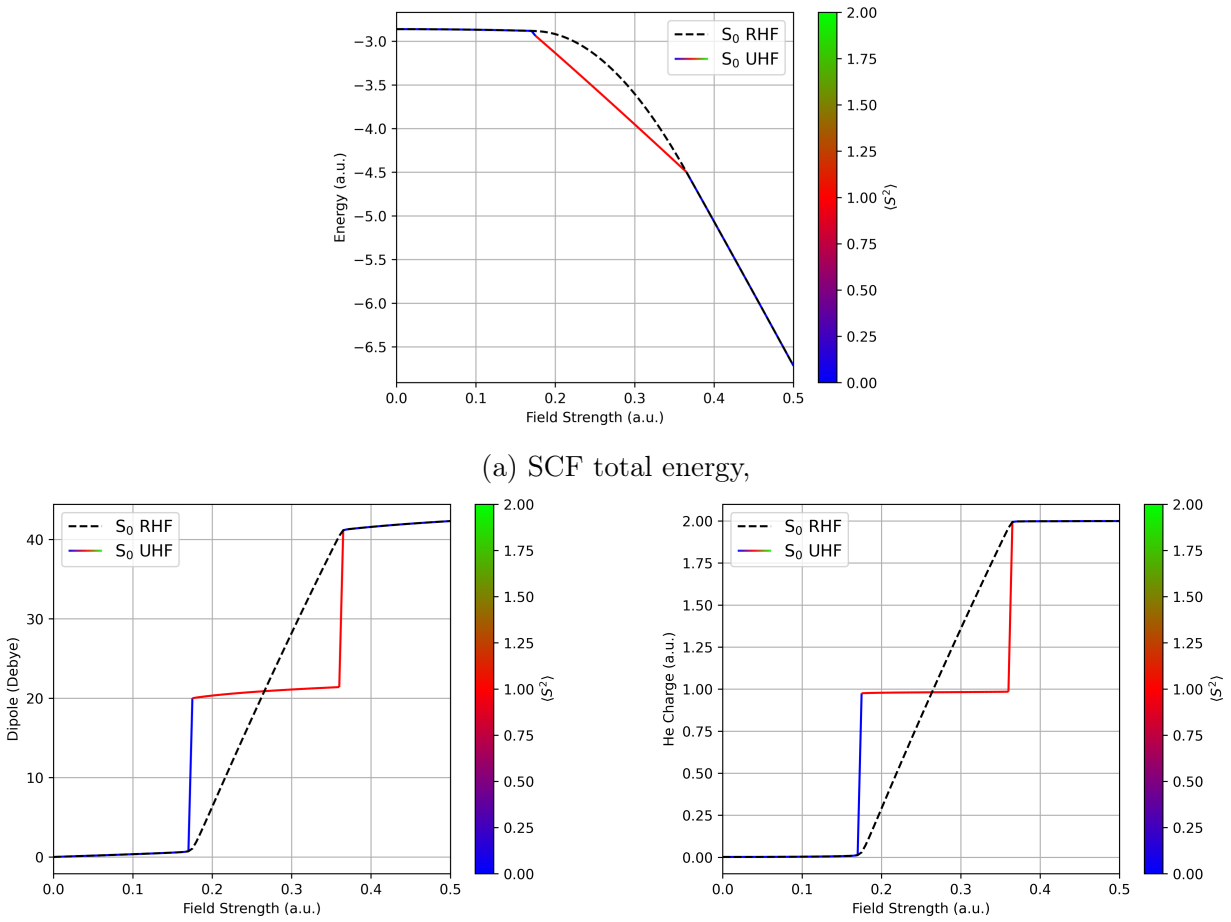
The analysis of the unrestricted case is more intricate and interesting. Two independent variational parameters (θ_1, θ_2) control the mixing of the basis functions for the independent sets of α and β orbitals (Table 2.1). The stationary condition on the energy expectation value for the UHF determinant leads to four different solutions for the (θ_1, θ_2) pairs. The solutions $(\theta_1, \theta_2) = (0, 0)$ and $(\theta_1, \theta_2) = (\frac{\pi}{2}, \frac{\pi}{2})$ reduce to RHF, corresponding to double occupation of the atomic and ghost orbitals, respectively. The former case is stable when $Fd < IP_1$, whereas the stability for the latter is achieved when $Fd > IP_2$. The difference in the UHF case is the possibility of achieving intermediate stable solutions $(\theta_1, \theta_2) = (0, \frac{\pi}{2})$ and $(\theta_1, \theta_2) = (\frac{\pi}{2}, 0)$ for the intermediate range of field strengths $IP_1 < Fd < IP_2$. These extra UHF solutions are characterized by having one electron occupying the atomic orbital and the other one localized on the ghost function. Therefore, we see that UHF is able to characterize the singly ionized state. This ability, however, comes at a cost: in an analogy to the common problem of bond dissociation at the UHF level, these single determinant unrestricted solutions for the intermediate range of field strengths are composed of an equal

mixing of singlet and triplet configurations that results in a spin-contaminated state with $\langle S^2 \rangle = 1$[24]. Nonetheless, UHF is qualitatively able to describe the static field ionization behavior expected from our analysis of the exact results (Fig. 2.2), showing the three distinct regimes corresponding to the three possible levels of ionization. In addition, the large separation between $|A\rangle$ and $|C\rangle$ leads to a vanishing singlet-triplet gap, so the contamination has no direct impact on the energetics. Finally, we note that the 1:1 quantitative mapping between the position of the "kinks" in the energy curve as a function of field strength and ionization potentials (IPs) is only possible if the conditions outlined in Table 2.1 are satisfied. The model can be easily modified to account for the electrostatic repulsion between $|A\rangle$ and $|C\rangle$ and the kinetic energy of the populated ghost function ($\langle C|\hat{T}|C\rangle$, leading to a better numerical agreement for the IPs. Alternatively, a decomposition of the total energy into atomic, continuum and interaction contributions could lead to better quantitative agreement for the IPs. Qualitatively, however, the main features do not depend on the specific nature of the ghost functions.

2.3.2 Wavefunction Methods in a Larger Basis

We move on to an analysis of the performance of other wavefunction based methods on a larger basis set to describe the static field ionization of He. We should point out that our emphasis is, once again, on the difference between how different flavors of self-consistent field (SCF) solutions can capture ionization effects in this discretized continnum model. In this case, He is represented by the larger aug-cc-pVTZ basis and the continuum is represented by a track of 18 1s hydrogenic STO-3G ghost basis functions spanning the range between -4.5 Å and 4.5 Å (with equal spacing of 0.5 Å between each ghost center) along the direction of the field. Fig. 2.3a summarizes the results in this larger basis set.

Just as expected from the analytical model, RHF cannot qualitatively describe the correct behavior for single-electron ionization due to the constraint that the two electrons are paired. For field strengths smaller than 0.175 a.u., we observe the neutral (two electrons localized around the atom) ground state, and for field strengths greater than 0.36 a.u., we observe the doubly ionized state of the atom. We notice a smooth transition between neutral atom and its atom dication state for intermediate field strengths. This can also be observed when we analyze how the total electric dipole (Fig. 2.3b) of the system and the charge of the He atom (Fig. 2.3c) change as a function of the applied field. UHF, on the other hand, captures the essential features expected from the exact solution: the energy curve shows three distinct regions separated by first derivative discontinuities. These kinks are associated with spin-polarization effects as illustrated by the transition between $\langle S^2 \rangle = 0$ to $\langle S^2 \rangle = 1$. The sharpness of this transition (*i.e.*, how fast the spin polarization occurs) is highly dependent on the size of the basis set: for smaller basis sets, $\langle S^2 \rangle$ seems to continuously change from 0 to 1, which indicates another remarkable similarity between the analysis presented in this work and the nature of the Coulson-Fischer point for bond dissociation. For larger basis sets, the spin polarization transitions appears to be much more sudden and discontinuous. We next consider the behavior of electric dipole and He charge for the UHF solution (Figs. 2.3b



(a) SCF total energy,

(b) Dipole moment,

(c) Mulliken charges on He,

Figure 2.3: (a) SCF total energy, (b) Dipole moment and (c) Mulliken charges on He as a function of field strength for the RHF and UHF solutions. In all cases, the RHF curve smoothly changes as the field strength increases, while UHF presents the characteristic kinks that are associated to ionization events.

and 2.3c). We do however note that the numerical magnitude of the change in the dipole moment is directly proportional to the size of the box created by the ghost functions. Both observables indicate that UHF qualitatively captures the three distinct regimes expected for the ionization of helium: the polarization of the electron pair on the neutral atom, the singly ionized state and the doubly ionized state. Again, the ability of UHF to describe this singly ionized state comes at the cost of spin polarization as indicated by $\langle S^2 \rangle = 1$ for field strengths in the range that generates one unit of positive charge on helium.

Moreover, we briefly analyze the MP2 (Fig. 2.4) performance for strong field ionziation of He using restricted (RMP2) and unrestricted (UMP2) orbitals. As expected, since UHF

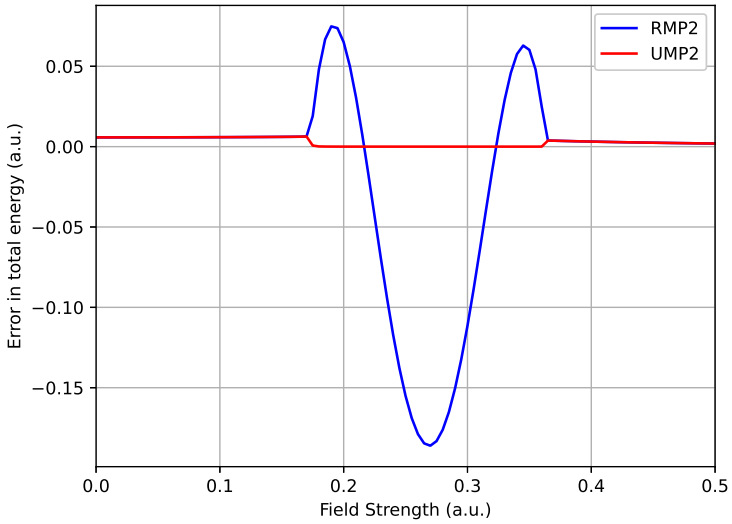


Figure 2.4: RMP2 and UMP2 results for the static field ionization of He. Due to the poor performance of the RHF reference, RMP2 presents energies lower than the exact ones, indicating that the MP2 amplitudes are overcompensating for the bad HF reference in the singly ionized regime. On the other hand, UMP2 is able to correctly recover almost exactly the FCI.

is a better reference to describe all possible charge regimes for our systems, UMP2 recovers all of the qualitative features of the exact solution. By contrast, RMP2 seems to oscillate above and below the exact solution: for field strengths closer to the ionization points, RMP2 energies are, as expected, above the exact values and seem, at first, to follow the same pattern as the RHF curve. However, RMP2 drops below the exact curve for fields in the mid-range for the singly ionized state. This indicates the RMP2 is over-correcting for the bad reference provided by the restricted orbitals in this regime. Nonetheless, this behavior is not as catastrophic as in the RMP2 potential energy curve for the dissociation of H_2, where incipient orbital degeneracies cause a divergent PES as the bond length increases [24]. Although not explored in the present work, this failure could be a fertile ground for an investigation of the efficiency of orbital-optimized MP2 schemes (OO-MP2) with regularization [136] to cheaply but accurately recapture some of the correlation effects associated to field ionization.

Our final analysis for the He atom is based on an attempt to remove most of the electric field and spurious discretized continuum contributions to the total energy, accounting only for the purely electronic energy of the atom as a function of the field strength. Within the dipole approximation, we defined an internal energy operator $\hat{U}$:

$$\hat{U} = \hat{H}_{elec} - \hat{\vec{\mu}} \cdot \vec{F} \tag{2.1}$$

The electronic energy of the combined atom-continuum system can then be calculated via:

$$E_{elec} = \langle \hat{H}_{elec} \rangle = E_{elec}^{atom} + E_{elst} + E_{elec}^{ghost} = \langle \hat{U} \rangle + \langle \hat{\vec{\mu}} \rangle \cdot \vec{F} \tag{2.2}$$

Eq. 2.2 includes three main components: (i) the electronic energy of the isolated atom (E_{elec}^{atom}), (ii) the electrostatic interaction between discretized continuum and He atom (E_{elst}), and (iii) the electronic energy of the populated discretized continuum (E_{elec}^{ghost}).

In the limit of two non-interacting subsystems (the isolated atom and the discretized continuum represented by the ghost functions), the super-system wavefunction is given by the product of the two individual wavefunctions and the super-system density matrix (D) can be written as a direct sum of the atomic density matrix (D^A) and a density matrix associated with the ghost functions (D^G). For our setup, however, the definition of the atomic subsystem is somewhat fuzzy since, in the complete basis set limit for the atom, the most diffuse basis functions will overlap with the nearest ghost centers added to represent the continuum electrons. This situation could be ameliorated by increasing the distance between this nearest ghost center and the atom or by noticing that, after ionization, electrons will tend to occupy the ghost functions closer to the edge of the artificial box used for the discretization of the continuum states. This allows us to have a clear definition of the ghost subsystem and we can then define the atomic contribution as the difference between the total energy and the energy contributions from the ghost basis functions.

By following this approach, D^A and D^G are defined as blocks of the super-system density matrix associated with the electron density represented by basis functions localized around the atom and with the density represented by basis functions centered in the track of ghost centers, respectively. The electrostatic interaction can then be calculated as

$$E_{elst} = \sum_{\mu,\nu \in G} D^G_{\mu\nu} V^G_{\mu\nu} + \sum_{\mu,\nu \in G} \sum_{\lambda,\sigma \in A} D^G_{\mu\nu} (\mu\nu|\lambda\sigma) D^A_{\sigma\lambda} \tag{2.3}$$

In Eq 2.3, we separated the contribution of the attractive interaction between the electrons populating a ghost function and the He nucleus (which is accounted by the ghost function block of the nuclear attraction one-electron integrals, V^G) and the repulsive screening between the continuum-like electron density (D^G) and the remaining electron density localized around He (D^A).

The electronic energy of the populated discretized continuum is expressed as a trace over its subsystem density matrix, D^G:

$$E_{elec}^{ghost} = \frac{1}{2} \sum_{\mu,\nu \in G} (F^G_{\mu\nu} + T^G_{\mu\nu}) D^G_{\mu\nu} \tag{2.4}$$

where T^G is the kinetic component of the one-electron integrals for the ghost basis functions, F^G is a Fock matrix in which we replaced the usual one-electron integrals by only its kinetic component (since for the discretized continuum subsystem there is no nuclear attraction).

Eq. 2.4 is clearly correct in the non-overlapping limit. In the presence of overlap, Eq. 2.4 corresponds to the embedded mean-field theory [221, 222] decomposition of the total energy of two composite systems (in our case He and ghost centers) into the energies of each subsystem and their interaction energy. It is worth noticing that we have not explicitly included other energy contributions arising from the mixed atom-ghost density contributions

or atom-ghost exchange effects in Eq. 2.2. Implicitly, these contributions are accounted for in the definition of E_{elec}^{atom} (Eq. 2.5) and are small due to large separation between the atom and the ghost function on the edge of the track.

Finally, the quantity of interest ΔE_{elec}^{atom} is given in terms of the atomic subsystem electronic energy at a given field strength

$$E_{elec}^{atom}(F) = \langle \hat{U} \rangle + \langle \hat{\vec{\mu}} \rangle \cdot \vec{F} - E_{elst} - E_{elec}^{ghost} \tag{2.5}$$

relative to the corresponding quantity at zero field:

$$\Delta E_{elec}^{atom} = E_{elec}^{atom}(F) - E_{elec}^{atom}(F = 0) \tag{2.6}$$

Based on our previous analysis, this approach should be more suitable for the UHF solution in which the separable density assumption holds to a greater extent and as it can properly account for ionization. Fig. 2.5 shows that the kinks in the total UHF energy (Fig. 2.3a) and the dipole and charge discontinuities for the UHF solution (Figs. 2.3b and 2.3c) are, once again, intrinsically related to ionization events, as we can directly recover information about the first and second ionization potentials (represented by the dashed lines in Fig. 2.5) of He by removing the spurious field and ghost functions contributions to the total energy. Finally, the analysis based on the decomposition of the density matrix into its atomic and ghost contributions allows us to obtain more information about the electronic structure of each subsystem: an eigenvalue decomposition of the combined atom-ghost, atomic and ghost density matrices indicates the number of electrons for each subsystem as illustrated in Fig. 2.6.

2.4 Field Ionization of a neon atom

In the previous section we thoroughly discussed how spin polarization is essential to qualitatively describe ionization by static fields in our model in which the continuum was discretized by a track of ghost functions placed along the direction of the applied electric field. At least for the He atom, we concluded that UHF properly captures, at the expense of spin contamination, the features associated with ionization within this limited model: we observe kinks in the potential energy curves as a function of field strength that are accompanied by discontinuities in the total electric dipole moment and the charge on the He atom. Now we move on to investigate if this simplest flavor of spin polarization (*i.e.*, allowing α and β orbitals to have different spatial parts) is enough to capture the features of atomic ionization in a heavier atom. Fig. 2.7a shows the RHF and UHF energy for the Ne atom. Once again, we see characteristic spin polarization and kinks in the UHF potential energy surface that are associated with ionization. We performed the same previous analysis for the electric dipole moment (Fig. 2.7b) and Mulliken charges on the Ne atom (Fig. 2.7c) and they both agree with our previous discussion: at the expense of losing information about the total spin of the system (*i.e.* spin contamination and our wavefunction not being an eigenstate of the

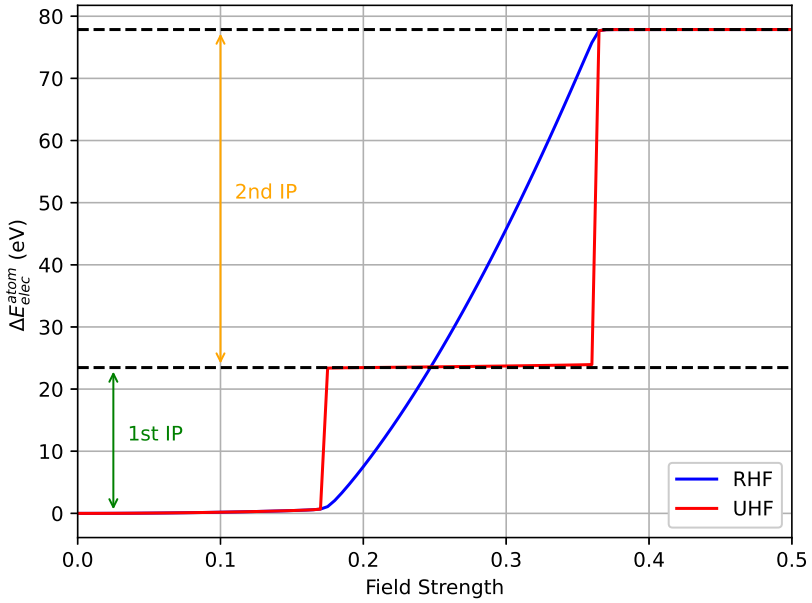
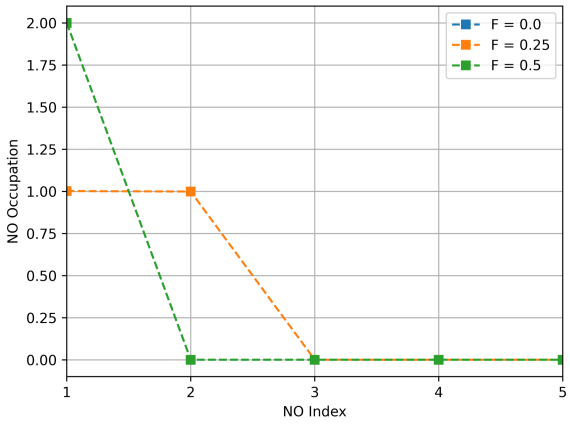


Figure 2.5: ΔE_{elec}^{atom} as defined in Eq 2.6. UHF is able to correctly recover information about both first and second ionization potentials of He, whereas RHF, missing a proper description of the first ionization event, only captures the second IP. The IP's were calculated at the HF/aug-cc-pVTZ+ghost track level of theory.

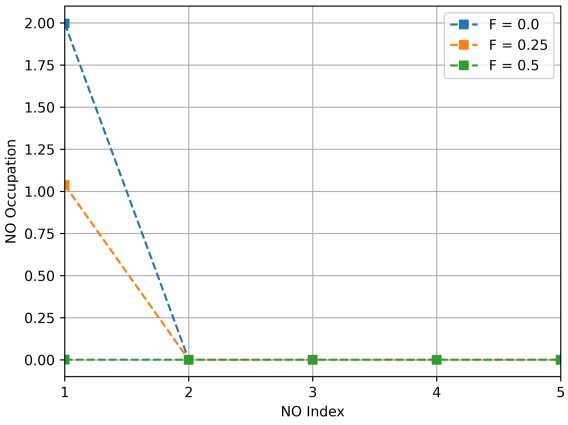
$\hat{S}^2$ operator), we can now ionize an odd number of electrons for Ne through UHF, whereas those states are not accessible through a RHF reference.

The electronic structure of Ne is more complicated than He which also allows us to analyze other qualitative features of the ionized states of the system as the strength of the applied field increases. With this is mind, it is important to point out what one would expect for each of the Ne ionized states within this limited discretized continuum model that we used in our analysis. Table 2.2 summarizes simplified electronic configurations expected for Ne, Ne^+, Ne^{2+} and Ne^{3+} and the track of ghost functions (with net charge $0, -1, -2, -3$) that represent the continuum-like states. Even though a proper analysis of the nature of the real ionized states of the atom is hindered due to the single determinant nature of HF theory, one can notice that the S_1 and S_2 states in Table 2.2 corresponds to one of the components of the physical (field-free) ^2P state of the Ne cation and ^3P state of the Ne dication, respectively. Our goal now is to see how we can obtain qualitative information about these physical states through Hartree-Fock theory within our limited model for ionization.

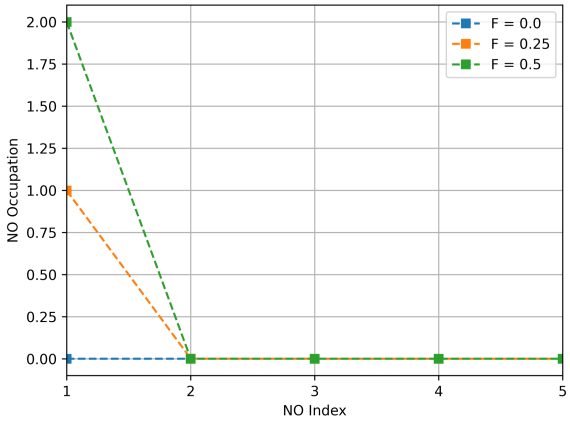
For the first ionization, we expect one unpaired electron at both the atom and the ghost functions, leading to a spin-contaminated UHF state in a similar fashion to what was described for the He case (Fig. 2.6). Fig. 2.8 summarizes some of the features of the UHF for a field strength of $F = 0.25$ a.u., which lies within the region for single ionization. We can



(a) NOONs for He + ghost system



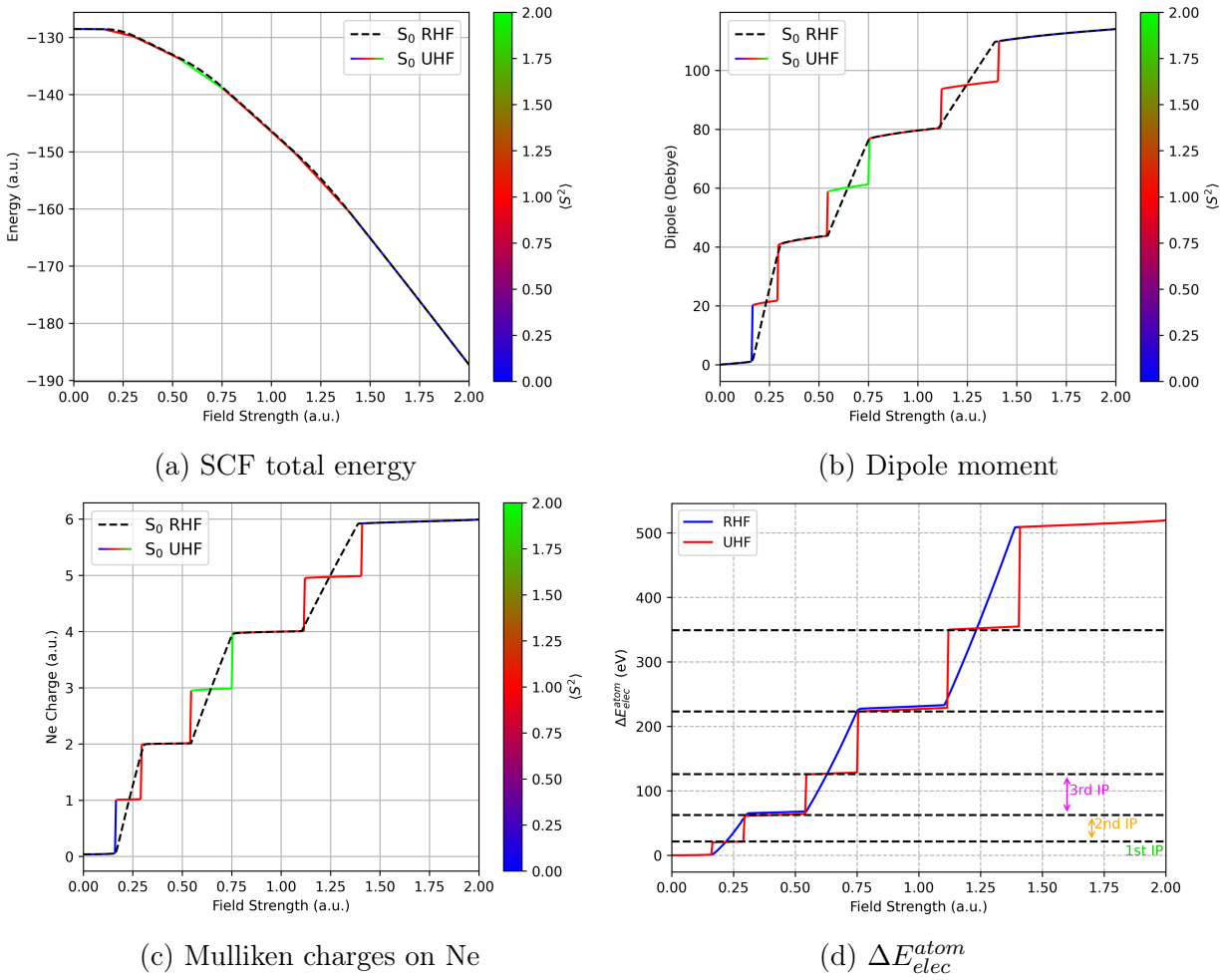
(b) NOONs for He subsystem



(c) NOONs for ghost subsystem

(d) Spin density for F = 0.25 a.u.

Figure 2.6: Natural Orbital Occupation Numbers (NOONs) for (a) He + Ghost complex system, (b) He subsystem (c) Ghost subsystem for the ionization of He in different regimes: neutral (F = 0 a.u.), singly ionized (F = 0.25 a.u.) and doubly ionized (F = 0.5 a.u.). (d) Spin density for the singly ionized state of He.



(a) SCF total energy

(b) Dipole moment

(c) Mulliken charges on Ne

(d) ΔE_{elec}^{atom}

Figure 2.7: Hartree-Fock calculations of the field ionization of Ne using the aug-cc-pVTZ basis set augmented by hydrogenic 1s STO-3G functions at the ghost centers. (a) SCF total energy, (b) Dipole moment, (c) Mulliken charges on Ne and (d) ΔE_{elec}^{atom} as defined in Eq 2.6 as a function of field strength for the RHF and UHF solutions. In all cases, the RHF curve smoothly changes as the field strength increases, while UHF presents the characteristic kinks that are associated with ionization events. The IP's were calculated at the HF/aug-cc-pVTZ+ghost track level of theory.

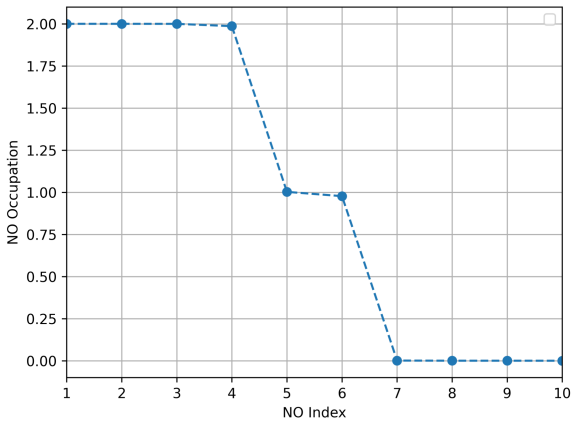
identify two unpaired electrons for the whole system (Fig. 2.8a). Performing an eigenvalue analysis on the partitioned density matrices for the atom subsystem (D^A) and the continuum/ghost subsystem(D^G) to obtain the natural orbital occupation numbers (NOONs) reveals that, as expected, each subsystem has a single unpaired electron (Figs. 2.8b and 2.8c): an α electron for the discretized continuum and an excess β electron for the atom (Fig. 2.8d), keeping $M_s = 0$ for the contaminated ground state.

Ionization State	Ne configuration (only $2p$ orbitals)	Ghost configuration
$\|S_0\rangle$ (neutral)	$\|2p_x^2 2p_y^2 2p_z^2\rangle$	$\|0\rangle$
$\|S_1\rangle$ (singly)	$\|2p_x^2 2p_y^2 2p_z^\alpha\rangle$	$\|C_1^\alpha\rangle$
$\|S_2\rangle$ (doubly)	$\|2p_x^2 2p_y^\alpha 2p_z^\alpha\rangle$	$\|C_1^2\rangle$
$\|S_3\rangle$ (triply)	$\|2p_x^\alpha 2p_y^\alpha 2p_z^\alpha\rangle$	$\|C_1^2 C_2^\alpha\rangle$

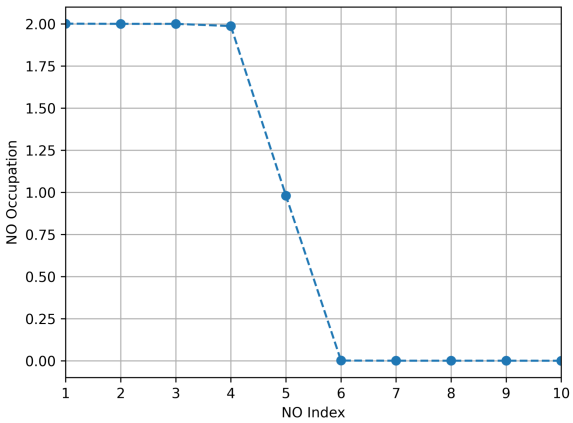
Table 2.2: Simplified electronic configuration of Ne and ghost basis functions for the neutral atom and the first 3 ionized states $|\mathrm{S}_i\rangle$. Each C_i represents a ghost orbital that can be populated by ionized electrons.

The doubly ionized state in He simply corresponds to the pairing of the two electrons in one of the ghost functions, leaving behind just the atomic nucleus (Fig. 2.6, the curve indicated by a field strength of F = 0.5 a.u.). For Ne, however, we expect the second ionized electron to pair with the first one and localize around a ghost center while Ne becomes a triplet state characterized by two unpaired electrons with the same spin and consequently total $M_s = \pm 1$ (see Table 2.2). This state, however, cannot be achieved by following the UHF potential energy curve which is constrained to the $M_s = 0$ subspace. Instead, the UHF solution for this doubly ionized state of Ne is characterized by a $\langle S^2 \rangle = 1$ spin contaminated state comprised of an α electron and a β unpaired electrons on the atom and no unpaired electrons for the track of ghost functions, as illustrated in Fig. 2.9. Some physical insight can be gained here by analyzing which states of the field-free Ne dication are included in the UHF $M_s = 0$ field dependent solution in the doubly ionized regime: the external field, breaking spherical symmetry, mixes the ^1S and ^1D states of Ne^{2+}. Moreover, due to spin symmetry-breaking in UHF, we observe that one of the components ($M_s = 0$) of the ^3P is also included in the mixture. It should be noticed that such behavior also highlights one of the limitations of our approach to describe the ionized flux: by constraining electron density to escape the atom through a unique track of ghost functions, we observe a driving force towards pairing of the ionized electrons. This prevents neon from achieving the desired and expected triplet configuration. A possible, yet not explored in this work, way to minimize this issue would be to expand our description of the discretized continuum by adding more ghost functions into two separate parallel tracks (as illustrated by the grey crosses in Fig. 2.1), which would possibly allow the delocalization of the ionized electrons, leaving the atom subsystem with $M_s = \pm 1$ and the ghost subsystem with $M_s \mp 1$, while maintaining the total $M_s = 0$ for the UHF solution.

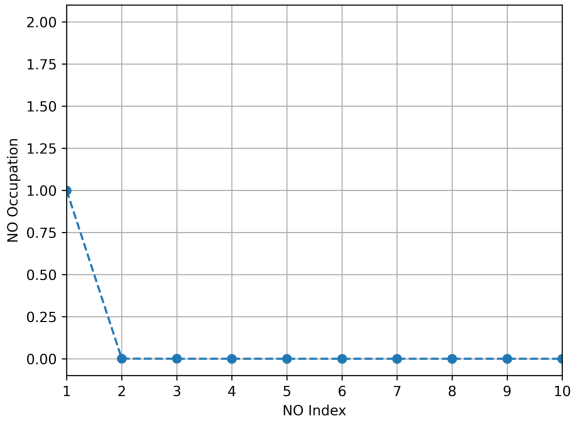
Alternatively, since the main interest should be to capture as much physical insight for the atomic subsystem as possible within our limited model, it is possible to remove the $M_s = 0$ constraint in UHF by exploring the manifold of generalized Hartree-Fock solutions (GHF). In GHF, all of the exact spin symmetries of the system (namely $\hat{S}^2$ and $\hat{S}_z$) can be broken by taking spin-orbitals that no longer belong to two independent sets of α and β orbitals, but now are linear combinations of orbitals in both sets. This allows the spin-quantization



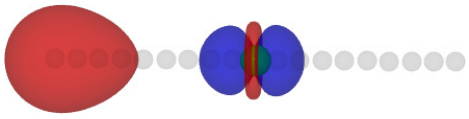
(a) NOONs for Ne + ghost complex system



(b) NOONs for Ne subsystem



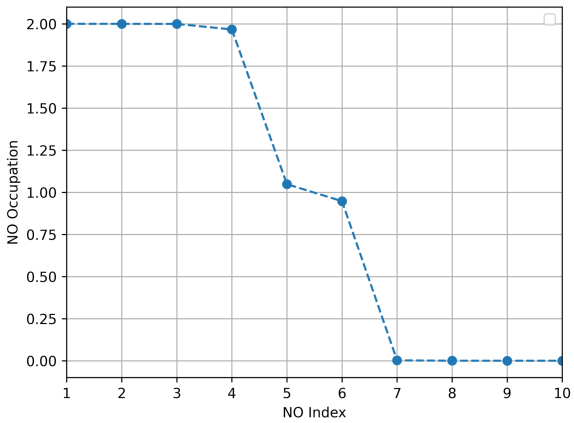
(c) NOONs for ghost subsystem



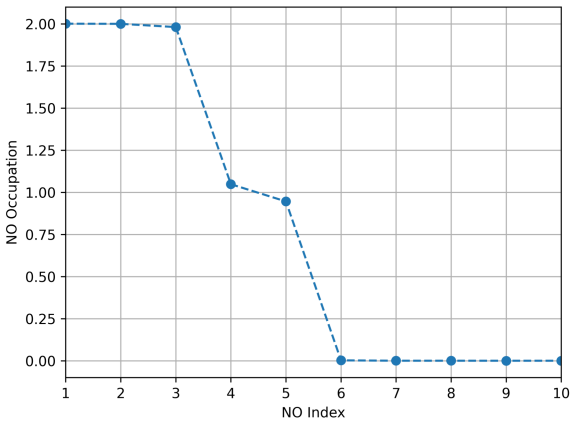
(d) UHF spin density

Figure 2.8: Analysis of UHF for the singly ionized regime of Ne (F = 0.25 a.u.), using NOONs for (a) Ne + Ghost complex system, (b) Ne subsystem (c) Ghost subsystem; (d) UHF spin density plotted for an isovalue of 0.002 a.u. (blue and red parts indicate an excess of α and β electrons respectively).

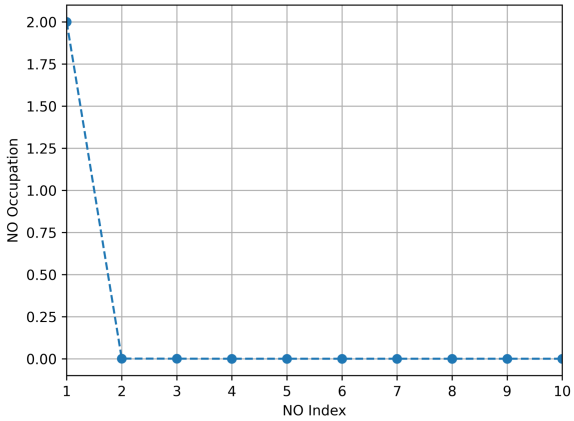
axis of individual electrons to rotate and not necessarily be parallel to each other in order to achieve the variationally lowest possible energy state. This extra flexibility does not mean that the GHF solution cannot be obtained through UHF though. A GHF wavefunction is said to be non-collinear (and sometimes referred to as a true GHF solution) if its spin cannot be quantized along some axis and this illustrates a real UHF $\rightarrow$ GHF instability of the orbital hessian [223, 224]. This is the case for the double bond dissociation in CO_2 [223, 225] and for systems presenting spin-frustration [226, 227]. If such a spin quantization axis exists, the GHF solution can be found by exploring UHF wavefunctions with different spin multiplicity or lying within different M_s manifolds. One can then find the difference between



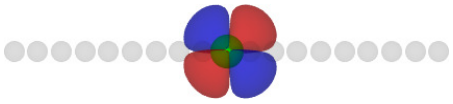
(a) NOONs for Ne + ghost complex system



(b) NOONs for Ne subsystem



(c) NOONs for ghost subsystem



(d) UHF spin density,

Figure 2.9: Analysis of UHF for the doubly ionized regime of Ne (F = 0.5 a.u.), using Natural Orbital Occupation Numbers (NOONs) for (a) Ne + Ghost complex system, (b) Ne subsystem (c) Ghost subsystem; (d) UHF spin density plotted for an isovalue of 0.002 a.u. (blue and red parts indicate an excess of α and β electrons respectively).

the number of up and down electrons by projecting the spin operators in this quantized axis [223]. In this context, we see that GHF can provide a convenient way to adiabatically follow the lowest UHF solution independently of M_s constraints.

Figure 2.10a illustrates the behavior of the different M_s UHF solutions compared to GHF for the strong field ionization of the neon atom within the race track of ghost functions to represent discretized continuum states that we have been using so far. Even though we cannot notice significant differences between UHF and GHF in terms of their description of charges, dipoles and total number of unpaired electrons for the supersystem, the GHF potential energy curve follows the lowest UHF solution, highlighting its collinear nature. For instance,

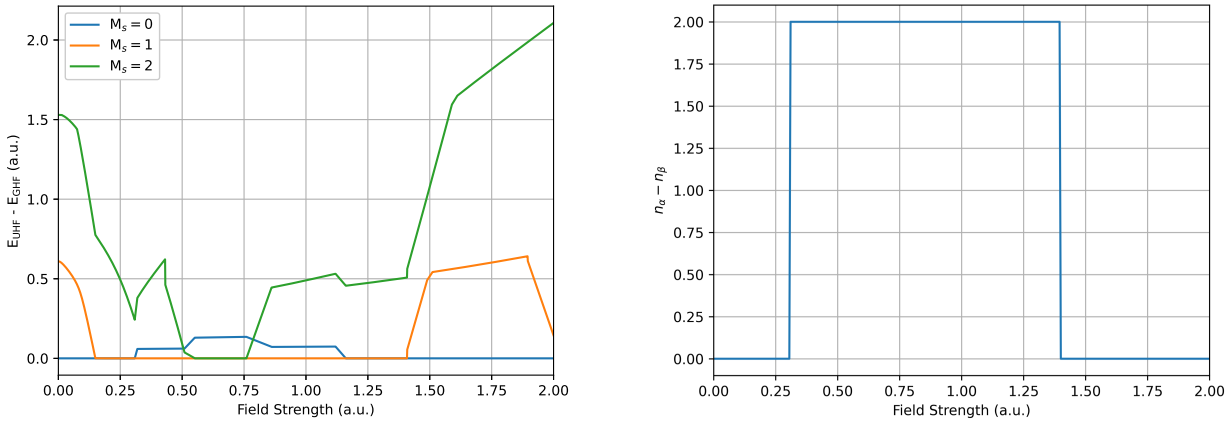
for field strengths lower than F = 0.15 a.u., we observe that the $M_s = 0$ UHF solution is the one with the lowest energy and that there is no difference between this solution and GHF, whereas the other M_s states are higher in energy. For electric fields ranging from F = 0.15 a.u. and F = 0.309 a.u. (single ionization), however, the $M_s = 0$ and $M_s = 1$ states become degenerate, characterizing the RHF→UHF instability previously discussed. For the regime between F = 0.309 a.u. and F = 0.60 a.u (which corresponds to the doubly ionized regime) the $M_s = 1$ UHF solution becomes the lowest in energy, illustrating the transition between different M_s subspaces that is captured naturally through GHF. Our first UHF analysis was constrained within the $M_s = 0$ subspace and, therefore, the wrong qualitative behavior (one α and one β electron) was observed for the double ionized state of the atom. GHF however, abolishes the $M_s = 0$ constraint and we can access the relevant doubly ionized state lying in the $M_s = \pm 1$ manifold (Fig. 2.10b) and Ne achieves the expected configuration with two unpaired α electrons which better resembles the $M_s = \pm 1$ component of the physical ^3P field-free state of the Ne dication.

Finally, even though the plateau regions in Figs. 2.7b and 2.7c are well characterized by collinear GHF solutions that represent UHF wavefunctions with different M_s values, the transition between these ionization states might not present the same behavior. At the onset of ionization, we suspect that, similarly to what was observed for the RHF → UHF instability, one would need a non-collinear GHF solution to properly describe the transition between one ionized state to the other. However, due to the dependence of the sharpness of this transition on the size of the artificial box delimited by the ghost functions, we did not pursue further investigation of this question. Hence, we note that GHF was used merely as a tool to access the lowest UHF state possible without any initial constraints on the number of α and β electrons.

2.5 Conclusions and Outlook

In conclusion, the present work explored an analogy between the Coulson-Fischer point that has been extensively characterized for bond dissociation problems and static strong field ionization phenomena. Such extension is, at first, hindered by the fact that ionization implies an outgoing electron flux that is poorly described by the atom-centered Gaussian basis sets commonly used in electronic structure packages [220]. Hence, our basic "race-track" model was introduced to account for this ionized flux by adding ghost basis functions spanning an appropriate region of space to allow some characterization of the continuum-like states of the ionized system.

A preliminary analytical analysis of the ionization problem of He in a minimum basis with a single distant ghost function hinted that ionization events, which would be characterized by the localization of the electron density at one of the auxiliary ghost centers of the "race-track", would not be well characterized by a restricted Hartree-Fock approach. In particular, the RHF solution could describe the neutral and the doubly ionized states of the system, but would lack any information about the singly ionized state given its electron-pairing



(a) Potential energy surfaces

(b) Supersystem GHF effective number of up and down electrons

Figure 2.10: (a) Potential energy curves as a function of field strength for Ne. It is possible to notice that, depending of the field strength, different M_s UHF solutions become the lowest in energy and GHF connects these different subspaces by variationally targeting the unconstrained lowest energy state. (b) Difference between number of up and down electrons along the spin quantization axis for the GHF solution, indicating that for the doubly ionized regime (between F = 0.309 a.u. and F = 0.60 a.u), we have two unpaired α electrons left on the Ne dication.

constraint. This analysis also indicated the existence of a RHF $\rightarrow$ UHF instability for the singly ionized range of field strengths that was akin to the triplet instability that leads to spin polarization in the bond dissociation problem. The predictions from this toy model were validated by computations in a larger basis set. Moreover, the UHF potential energy surface as a function of field strength reveals kinks that, after an analysis of the charge population of the atom, the dipole moment of the system and the expectation value of the $\hat{S}^2$ operator, indicate the existence of an analog of the Coulson-Fischer point for static strong field ionization. We have also made an effort to eliminate possible spurious contributions stemming from our limited model of the discretized space by devising a partitioning scheme based on taking traces of the RHF and UHF density matrices over the appropriate basis functions centered only around the atom. Such scheme allowed us to, once again, determine the character of each state of the atom for different field strengths, as well as to recover information about the ionization energies of He. While RHF cannot represent the singly ionized state, UHF recovers the first and second ionization energies of He at the cost of breaking spin symmetry and being spin contaminated.

We then extended our analysis to a more complicated system: the Ne atom. The main takeaways are still valid: by constraining the HF solution to be an eigenfunction of $\hat{S}^2$, *i.e.* constraining ourselves to the RHF manifold of solutions, we cannot describe ionization of

an odd number of electrons as the strength of the external field increases. By exploring the solutions contained in the spin-polarized UHF manifold, on the other hand, we have a qualitatively correct description of the atomic charge associated with these ionized states. However, due to a strong drive to pair the electrons at the end of the auxiliary "race-track" of ghost functions and the constraint imposed on the value of $\langle \hat{S}_z \rangle$ for UHF solutions, the results do not correspond to the expected Hund's rule configuration of Ne for some field strengths. Thus, we also analyzed the use of GHF solutions to obtain a better qualitative description of the static strong field ionization of Ne. We observed that, while maintaining the same behavior as UHF for the atomic charges, GHF can properly rotate spins at the cost of losing information about the expectation value of $\hat{S}_z$ in order to give the proper high-spin description of the unpaired electrons on the atomic subsystem. Exploring how different kinds of symmetry-broken solutions, such as allowing complex polarization in Hartree-Fock (cRHF, cUHF and cGHF) [228–231], could also prove potentially useful to obtain qualitatively correct description of strong field ionization of more complex systems.

Even though the results and analysis presented in the current work were obtained for the case of an applied external static electric field, it can be argued that some of the consequences of imposing spin constraints on the Hartree-Fock solutions could also lead to a qualitatively wrong description of the real-time dynamics of the system when a physically meaningful laser pulse is used. For the more common problem of bond stretching in electronic structure, it has already been shown that using a qualitatively bad HF solution as the starting point for more elaborate methods that account for dynamic electron correlation could lead to failures in obtaining a good description of potential energy surface, for example. In this sense, work towards extending the analysis here to real-time time-dependent electronic structure is underway in our group. To achieve more physically meaningful results, a better approach to represent the continuum states, perhaps through a mixed Gaussian-plane wave basis set or complex absorbing potentials, is needed [185, 232].

Supplementary Material

Supplementary information for this work can be accessed via Ref. 233. They include raw data and additional plots.

Chapter 3

Changes in Polarization Dictate Necessary Approximations for Modeling Oscillator Strengths

3.1 Introduction

Investigating the electronic structure of materials is of paramount importance, not only to advance scientific understanding, but also from a technological perspective. Electronic structure is typically probed in terms of neutral or charged electronic excited states. A complete investigation of excited states involves the determination not only of the energies of the states but also of the transition probabilities for excitation (de-excitation) to (from) them. There are various experimental methods for this purpose, which provide opportunities to validate and advance theoretical efforts. For decades, density-functional theory (DFT) [55, 58] has been the primary workhorse in electronic structure calculations [234–237]. Even though DFT is a ground state theory, various techniques build upon DFT by leveraging different levels of approximation to probe electronic excited states [238–248]. The two main avenues for this, both of which are widely used in practice, are: the response-based treatment, where a response-function is evaluated indirectly by creating electron-hole (e-h) pairs in a reference state; and the orbital-constrained treatment, which models the target state using constrained non-aufbau orbital occupancy.

In this paper we link these two disparate approaches within a generalized framework. For valence-to-core electronic transitions, we show analytically, as well as with computed examples, that within the computationally tractable adiabatic approximation, inaccuracy in the response-based approach is correlated with the change in the valence polarization due to the core-ionization of the system. Additionally, working within the orbital-constrained approach, by modifying the state-of-the-art orbital-optimized DFT (ooDFT) formalism with physically-motivated approximations, we propose a new computational method for simulating non-resonant X-ray emission, with a focus on accurate calculation of the transition

amplitude.

Electronic transitions, which can involve both the core and the valence electrons, typically span a wide energy range. In order to ensure tractability of our calculation, we focus on electronic de-excitations since this allows us to limit our attention to the occupied subspace only, offering appreciable computational advantage. De-excitations can either be confined fully within the valence subspace, or involve valence-to-core transitions. Due to the availability of experimental data, the latter kind, which is experimentally accessible within the non-resonant X-ray emission spectroscopy (XES) technique, will be the primary focus of our computational demonstrations. Notably, since the goal of this work is the analysis of purely electronic de-excitations, additional effects which might influence the experimental spectra, e.g. the ionic vibrations present at finite temperature [249, 250] (note that vibrational effects typically make the spectra smoother by introducing additional broadening [251] and can occasionally introduce new peaks [252]) have not been taken into account.

3.2 Non-resonant X-ray emission formalism

In non-resonant XES [253–256], spontaneous X-ray emission (or fluorescence) occurs from an initially core-ionized sample. As the core-excited system decays from an initial state Ψ_I with energy E_I to any possible final state Ψ_F with energy E_F, the emission intensity

$$I(\omega) \propto \sum_F (E_I - E_F)^4 |M^{I,F}|^2 \delta(\omega + E_F - E_I), \tag{3.1}$$

is recorded as a function of its frequency ω, where, denoting the many-body transition operator by $\hat{O}$, the transition matrix element is given by $M^{I,F} = \left\langle \Psi_F \middle| \hat{O} \middle| \Psi_I \right\rangle$. Therefore $|\Psi_\mathrm{I}\rangle$, i.e., the state prior to the X-ray emission process, can be represented by the lowest energy state of the positively charged system with a core-hole on the excited atom. In our calculations, this is approximated by a single Slater determinant (SD) composed of Kohn-Sham (KS) orbitals obtained from a self-consistent field (SCF) DFT calculation, where one core orbital is constrained to be empty, which we denote as the zero-th orbital. One numerically stable approach to obtain this core-hole excited state directly is the maximum overlap method (MOM) [257]. If there are N electrons remaining in the system, then, denoting the creation operator for the j-th initial-state orbital $\tilde{\phi}_j$ by $\tilde{a}_j^\dagger$, we can represent the SD as

$$|\Psi_\mathrm{I}\rangle = \left(\prod_{j=1}^{N} \tilde{a}_j^\dagger \right) |0\rangle \,. \tag{3.2}$$

Assuming the core is filled in all constituent SDs, the exact F-th final state resulting

from de-excitation (filling the zero-th orbital via $\tilde{a}_0^\dagger$) can then be formally expressed as

$$\begin{aligned}\left|\Psi_F^{\text{exact}}\right\rangle = &\left[\sum_{j=1}^{N} \alpha_j^F \tilde{a}_j \tilde{a}_0^\dagger \left|\Psi_{\text{I}}\right\rangle\right] + \\ &\left[\sum_{l,m=1}^{N} \sum_{p=N+1}^{\text{all}} \beta_{l,m,p}^F \tilde{a}_l \tilde{a}_m \tilde{a}_p^\dagger \tilde{a}_0^\dagger \left|\Psi_{\text{I}}\right\rangle\right] + \ldots\end{aligned} \tag{3.3}$$

where α and β are expansion coefficients, and we use tildes to indicate that these initial state orbitals include a core hole. The terms inside the first (second) set of square brackets represent the terms obtained from $|\Psi_{\text{I}}\rangle$ by single (double) substitutions, i.e., by creating one (two) e-h pairs in $|\Psi_{\text{I}}\rangle$. The previously excited core orbital is indicated, as stated above, by subscript 0. Notably, beyond the singles term, i.e., in doubles and beyond, we start populating the unoccupied valence subspace of $|\Psi_{\text{I}}\rangle$.

However, only the singles can contribute to the transition-dipole matrix (for simplicity, we drop the I superscript on M as the initial state, which is the core-ionized state, will be the same for all de-excitations in XES)

$$M_{\text{exact}}^F = \left\langle \Psi_F^{\text{exact}} \middle| \hat{O} \middle| \Psi_{\text{I}} \right\rangle = \sum_{j=1}^{N} (\alpha_j^F)^* \left\langle \phi_0 \middle| \hat{o} \middle| \tilde{\phi}_j \right\rangle, \tag{3.4}$$

since for higher substitutions, the overlap with $\hat{O}\left|\Psi_{\text{I}}\right\rangle$ vanishes. Here, $\hat{O}$ ($\hat{o}$) denotes the many-body (single-particle) transition operator, defined explicitly below.

3.3 XES within adiabatic linear response

Now, in principle, response-based frameworks, such as linear response time-dependent DFT (LR-TDDFT) and the Bethe-Salpeter equation (BSE), evaluate the absorption/emission spectra exactly from the imaginary part of the density-density response function χ – the change in electron-density as a function of external potential. The complex problem of finding χ is mapped into that of diagonalizing a fictitious, typically non-Hermitian, two-particle Hamiltonian H^{2p}. Even though, strictly speaking, the response-based approaches do not offer any access to the wavefunctions, within the Tamm-Dancoff approximation [258], the final state is often approximated [259] by an eigenstate of H^{2p} and expressed as a linear combination of Slater determinants obtained by creating an e-h pair on the initial state $|\Psi_I\rangle$ (analogous to the configuration-interaction singles approach) as

$$\left|\Psi_F^{\text{Resp}}\right\rangle = \sum_{j=1}^{N} \gamma_j^F \tilde{a}_j a_0^\dagger \left|\Psi_{\text{I}}\right\rangle, \tag{3.5}$$

such that the corresponding transition amplitude

$$M^F_{\text{Resp}} = \left\langle \Psi^{\text{Resp}}_F \middle| \hat{O} \middle| \Psi_{\text{I}} \right\rangle = \sum_{j=1}^{N} (\gamma^F_j)^* \left\langle \phi_0 \middle| \hat{o} \middle| \tilde{\phi}_j \right\rangle \tag{3.6}$$

is identical to that of a single-particle de-excitation from the auxiliary orbital

$$\left| \phi^{\text{Resp}}_F \right\rangle = \sum_{j=1}^{N} \left(\gamma^F_j\right)^* \left| \tilde{\phi}_j \right\rangle . \tag{3.7}$$

In response-based approaches, the e-h interaction is encoded with the help of a kernel K, which, in its exact form, is time-dependent. However, in the standard, computationally viable adiabatic approximation [248], this dynamic effect is ignored. Within this approximation, the number of accessible final states equals the number of e-h pairs that can be created on the initial state and each approximate final state is normalized [248]. The normalization constraint means that, unless all contributions beyond the singles vanish identically in Eq. 3.3 (which is possible only if the occupied valence subspace of $|\Psi^{\text{exact}}_F\rangle$ has no overlap with the unoccupied subspace of $|\Psi_{\text{I}}\rangle$), the set of coefficients $\{\gamma^F\}$ must differ from $\{\alpha^F\}$, resulting invariably in the relation $M^F_{\text{Resp}} \neq M^F_{\text{exact}}$.

This leads to the crucial inference that if there is appreciable difference between the valence occupied subspaces of the ground and the core-ionized state – if the valence electron density is significantly rearranged due, in this case, to the annihilation of the core hole in XES – then, adiabatic response-based approaches will inevitably incur errors in the oscillator strength. By definition, this rearrangement of the valence electronic density defines a polarization response due to the perturbation, which is proportional to the spatially non-local electronic susceptibility.

Since the susceptibility of a given ground state naturally invokes coupling to the unoccupied orbital subspace, the change in the aforementioned response, i.e., the change in the valence polarization, serves as a useful metric for the validity of the adiabatic approximation in the response-based treatment. This metric can be particularly important in XES, since the annihilation of a localized core hole may easily result in a significant renormalization of the occupied orbitals, leading to a large change in polarization.

Notably, this argument does not preclude the exact time-dependent kernel from yielding the accurate oscillator strength, since, in this case, the requirement of normalization will not be applicable for $\left|\Psi^{\text{Resp}}_F\right\rangle$. In other words, the inadequacy is a consequence of the adiabatic approximation, not of the response-based approach, which, in principle, calculates χ exactly. Note that, such valence-to-core de-excitations can help the formulation of accurate frequency-dependent kernels [260] by providing experimental data for verification of their efficacy.

This analytical conclusion can be demonstrated using the molecular examples shown in Fig. 3.1, where the polarization response is simply the change in the electrostatic dipole moment of the molecule. We compare the measured and simulated XES spectra of two

small molecules (chloromethane [261] and phenol [262]), with calculations using DFT orbitals produced by the Q-Chem code [263] using the ωB97M-V exchange-correlation (xc)-functional [79] and the large aug-cc-pCVQZ basis-set [215, 264]. The energy convergence threshold was set to 10^{-8} a.u., and 10^{-14} a.u. was used as the threshold for screening two-electron integrals. The SG-2 standard quadrature grid is used for the DFT calculations which use the Pulay DIIS algorithm for convergence. A value of 10^{-8} a.u. is chosen as the threshold for the LR-TDDFT calculations.

Comparison with the experimental spectra reveals that the adiabatic LR-TDDFT spectrum is in good agreement for Cl $K\beta$ emission of CH_3Cl, which corresponds to a lower (1.43 D) change in the electrostatic dipole moment. However, for O $K\alpha$ emission of C_6H_5OH, the change is larger (3.01 D) and the mismatch with experiment is noticeable (particularly near 527 eV).

3.4 XES within the Orbital-constrained Approach

3.4.1 General Framework

Having characterized the shortcomings of the adiabatic response-based technique for simulating valence-to-core de-excitation intensity, we turn our attention to the orbital-constrained approach. Let us consider the general case of optical dipole transitions between many-body states, initial and final, comprising single-particle orbitals from distinct SCFs. The many-body transition operator can be written, with respect to the initial state, as:

$$\hat{O} = \sum_{i,j} \left\langle \tilde{\phi}_j \left| \hat{o} \right| \tilde{\phi}_i \right\rangle \tilde{a}_j^\dagger \tilde{a}_i, \tag{3.8}$$

where the ordering of operations is deliberate and the sum runs over the entire single-particle Hilbert space of the initial state. These transitions are individually weighted by their single-particle dipole matrix elements: $\tilde{o}_{ji} = \left\langle \tilde{\phi}_j \left| \hat{o} \right| \tilde{\phi}_i \right\rangle$, where $\hat{o} = \epsilon \cdot \hat{\mathbf{r}}$ defines the electric field polarization and its interaction with the electronic position. Whether these transitions are allowed or not depends on the many-body state and its orbital occupations. We will work within the single Slater determinant approximation, where the N occupied orbitals that define the self-consistent field via the electron density can also define a many-body state, emergent from the vacuum, as:

$$|\Psi\rangle = \prod_{i=1}^{N} a_i^\dagger |0\rangle, \tag{3.9}$$

where the creation and annihilation operators correspond to the orbitals of the relevant SCF. With this notation, we can define the many-body transition amplitude from an initial state

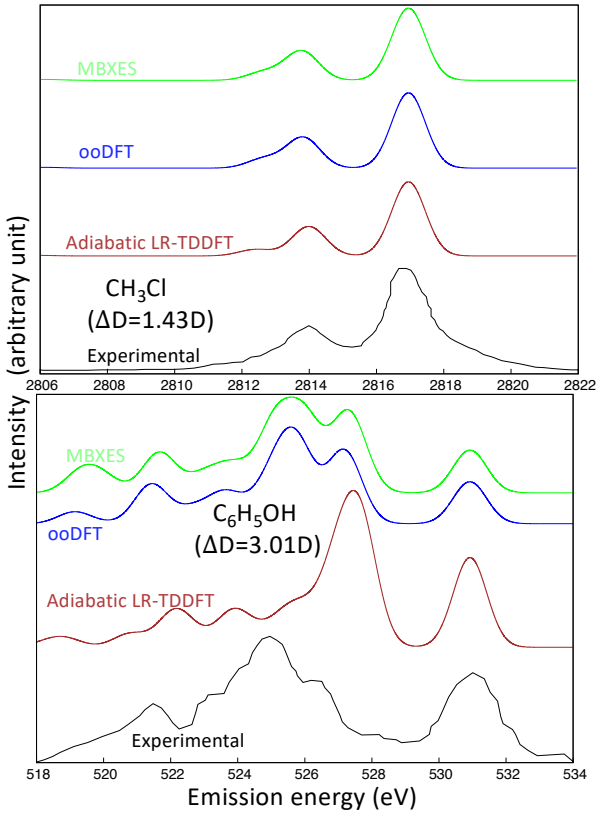


Figure 3.1: X-ray Emission spectra for (top) Cl $K\beta$ from chloromethane (CH_3Cl) and (bottom) O $K\alpha$ from phenol (C_6H_5OH) simulated using adiabatic LR-TDDFT (brown), ooDFT (blue), MBXES (green) and compared with experiment (black), from Refs [261] and [262], respectively. The associated change in dipole moment (ΔD) is indicated in each case. As detailed in the text, ΔD is a diagnostic for errors in LR-TDDFT arising from the adiabatic approximation, which is not required in ooDFT or MBXES (an approximation to ooDFT). A rigid shift is added to each simulated spectrum to align the highest-energy peak with the experiment.

$|\Psi_I\rangle$ to a final state $|\Psi_F\rangle$ as:

$$\begin{aligned} M^{I,F} &= \langle \Psi_F | O | \Psi_I \rangle \\ &= \sum_i^{\text{unocc}} \sum_j^{\text{occ}} \left\langle \Psi_F | \Psi_I^{+i-j} \right\rangle \left\langle \Psi_I^{+i-j} \middle| \hat{O} \middle| \Psi_I \right\rangle \\ &= \sum_i^{\text{unocc}} \sum_j^{\text{occ}} \left\langle \Psi_F | \Psi_I^{+i-j} \right\rangle \tilde{o}_{ij} \end{aligned} \tag{3.10}$$

where the functional form of O enforces that only single electron-hole pair excitations of the initial state can couple with the final state. The notation Ψ_I^{+i-j} implies the annihilation of the electron in orbital j and the creation of an electron in orbital i – a non-self-consistent creation of a single electron-hole pair within the orbitals defined by the initial state. Thus, $M^{I,F}$ can be expressed as a linear combination of purely single-particle transition amplitudes ($\tilde{o}_{ij}$) with the coefficients given by the many-body overlap terms $\left\langle \Psi_F | \Psi_I^{+i-j} \right\rangle$.

. (3.11)

${}^F\Xi_{+i-j}$ is essentially the determinant (with a factor of ± 1) of an $N \times N$ matrix composed of overlap terms between the occupied single-particle orbitals of the final and the initial state with the subscript $i - j$ indicating the removal (addition) of the j-th (i-th) electron in the initial state, which, for the aforementioned matrix, amounts to substitution of the j-th row by the i-th one.

For the XES process, Ψ_I is the state with a core-hole and can be approximated by the expression shown in Eq. 3.2. Consistently, we will employ the simplest case for the final state, Ψ_F, as representing a core-filled state with a single valence hole in orbital f, $|\Psi_F\rangle = \left|\Psi_F^{+0-f}\right\rangle$, where 0 denotes the core orbital. Due to the localized nature of the core electrons, it is reasonable to expect the core electron of Ψ_F to have negligible overlap with any electron present in Ψ_I, which lacks that core electron. Consequently, for XES, $\left\langle \Psi_F^{+0-f} | \Psi_I^{+i-j} \right\rangle \approx 0$ unless $i = 0$. This leads to the following simplification of Eq. 3.10

$$M^{I,F} = \sum_j^{\text{occ}} \left\langle \Psi_F^{+0-f} | \Psi_I^{+0-j} \right\rangle \tilde{o}_{0j}, \tag{3.12}$$

where, to reiterate, we have put a tilde over the transition-matrix element, to emphasize the fact that, for XES, the initial state contains a core-hole. We can expect the core-orbital subspaces of $\left|\Psi_F^{+0-f}\right\rangle$ and $\left|\Psi_I^{+0-j}\right\rangle$ to be approximately equal, for any f and j. First-order perturbation theory would support this approximation, especially due to the energy isolation of the core orbitals from the valence subspace – with perturbations scaling as the inverse of orbital energy differences. This was also observed numerically in our calculations. Therefore,

in the overlap term $\left\langle \Psi_F^{+0-f} | \Psi_I^{+0-j} \right\rangle$, the core-contribution, which can be approximated as unity, can be factored out and the overlap can be evaluated exclusively in terms of the valence orbitals. This is particularly useful in pseudopotential-based calculations, where only the valence KS orbitals are readily available.

For a final-state $|\Psi_F\rangle$ having a hole in the f-th orbital, the many-body overlap of Eq. 3.10 can now be written as

$$\left\langle \Psi_F^{+0-f} | \Psi_I^{+0-j} \right\rangle = {}^F C_{j,f} = \left[(-1)^{f+j}\right] {}^F m_{j,f} \tag{3.13}$$

such that

$$\begin{aligned} {}^F m_{j,f} = \\ \det \begin{bmatrix} {}^F\xi_{11} & \cdots & {}^F\xi_{1,f-1} & {}^F\xi_{1,f+1} & \cdots & {}^F\xi_{1,N} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ {}^F\xi_{j-1,1} & \cdots & {}^F\xi_{j-1,f-1} & {}^F\xi_{j-1,f+1} & \cdots & {}^F\xi_{j-1,N} \\ {}^F\xi_{j+1,1} & \cdots & {}^F\xi_{j+1,f-1} & {}^F\xi_{j+1,f+1} & \cdots & {}^F\xi_{j+1,N} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ {}^F\xi_{N,1} & \cdots & {}^F\xi_{N,f-1} & {}^F\xi_{N,f+1} & \cdots & {}^F\xi_{N,N}. \end{bmatrix} \end{aligned} \tag{3.14}$$

where

$${}^F\xi_{q,p} = \left\langle \phi_p^F | \tilde{\phi}_q^I \right\rangle \tag{3.15}$$

denotes the single-particle overlap between the p-th orbital of final-state F and the q-th orbital of initial-state I. Since the initial state is assumed fixed as the core-ionized state, we drop specific reference to state I, but we retain the superscript F as the single particle orbitals for each final state may vary, as we shall see below. From Eq. 3.13 and 3.14, ${}^{\mathbf{F}}\mathbf{C}$ and ${}^{\mathbf{F}}\mathbf{m}$ can be recognized respectively as the matrix of the cofactors and minors of ${}^{\mathbf{F}}\xi$, the $(N \times N)$ overlap matrix composed of the N lowest initial and final state valence orbitals

$${}^{\mathbf{F}}\xi = \{{}^F\xi_{q,p}\}_{q,p=1}^N. \tag{3.16}$$

We will make use of this realization later. The above expressions relied on two assumptions which we can reiterate here for clarity: (1) the many-body states are represented as single Slater determinants, which dictates that their overlap is also a determinant; (2) a specific core orbital 0 on a single atomic site is assumed unoccupied in the initial state and occupied in the final state.

3.4.2 Orbital-optimized DFT

Within the orbital-constrained framework, the orbital-optimized DFT (ooDFT) approximation employs a fully self-consistent procedure such that a final state with filled core and a hole in the f-th valence orbital is approximated by the SD:

$$|\Psi_F\rangle = \left|\Psi_{-f}^{(f)}\right\rangle = a_f^{(f)} \left(\prod_{j=1}^{N} \left(a_j^{(f)}\right)^\dagger\right) a_0^\dagger |0\rangle\,, \tag{3.17}$$

where $\left(a_j^{(f)}\right)^\dagger$ is the creation operator for the j-th KS orbital $\phi_j^{(f)}$ corresponding to the SCF of a system with a hole in the f-th orbital, preserved, for example, using MOM. There is a distinct set of operators for each final state F.

The resulting transition amplitude is then written as

$$M^F = M^f_{\text{ooDFT}} = \left\langle \Psi_{-f}^{(f)} \middle| \hat{O} \middle| \Psi_{\text{I}} \right\rangle = \sum_{j=1}^{N} C_{j,f}^{(f)} \tilde{o}_{0j} \tag{3.18}$$

where

$$C_{j,f}^{(f)} = \left\langle \Psi_{-f}^{(f)} | \Psi_{\text{I}}^{+0-j} \right\rangle \tag{3.19}$$

is the relevant determinantal overlap with $\left|\Psi_{\text{I}}^{+0-j}\right\rangle = \tilde{a}_j \tilde{a}_0^\dagger |\Psi_{\text{I}}\rangle$. Note that the transition moment is identical to that of the single-particle de-excitation of a non-interacting electron to the core level from an auxiliary orbital given by

$$\left|\phi_f^{\text{ooDFT}}\right\rangle = \sum_{j=1}^{N} C_{j,f}^{(f)} \left|\tilde{\phi}_j\right\rangle. \tag{3.20}$$

For phenol, the improvement in agreement with the experimental XES line shape is apparent in Fig. 3.1. In the SI [265], we provide a comparison of the range-separated and global hybrid functionals in Figs. S1 and S2 for phenol and in Figs. S3 and S4 for acetone.

3.4.3 The MBXES approach

A major drawback of the ooDFT method is that it requires a separate self-consistent field calculation for each final state of the system, rendering the computation formidably expensive, especially for large systems. Additionally, depending on the symmetry of the hole, certain configurations (especially near degeneracies) can be numerically difficult to converge.

To remedy this, we note that if the valence hole is sufficiently delocalized and, consequently, is less likely to induce a drastic change among the occupied electronic orbitals, we are justified in approximating ooDFT using a formalism which retains the self-consistent effect of the core-hole in the initial state but entirely neglects that of the valence-hole in the final state. For instance, a final state containing a hole in the f-th (say) valence orbital is approximated in terms of the $f$rozen ground state orbitals as

$$|\Psi_F\rangle = |\Psi_{-f}\rangle = a_f \left(\prod_{j=1}^{N} a_j^\dagger\right) a_0^\dagger |0\rangle\,, \tag{3.21}$$

where $a_j^\dagger$ is the creation operator for the $j-$th KS orbital (ϕ_j) of the ground state SCF. Defining the relevant determinantal overlap as $C_{j,f} = \left\langle \Psi_{-f} | \Psi_{\mathrm{I}}^{+0-j} \right\rangle$, the transition amplitude is then given by

$$M^f_{\mathrm{MBXES}} = \left\langle \Psi_{-f} \left| \hat{O} \right| \Psi_{\mathrm{I}} \right\rangle = \sum_{j=1}^{N} C_{j,f} \left\langle \tilde{\phi}_0 \left| \hat{o} \right| \tilde{\phi}_j \right\rangle \tag{3.22}$$

with

$$\left| \phi_f^{\mathrm{MBXES}} \right\rangle = \sum_{j=1}^{N} C_{j,f} \left| \tilde{\phi}_j \right\rangle \tag{3.23}$$

representing the corresponding auxiliary orbital. Note that since we now construct each $|\Psi_F\rangle$ from the frozen ground-state KS orbitals, the many-body overlap of Eq. 3.10, which in the ooDFT approximation is given by $C_{j,f}^{(f)}$ (Eq. 3.19), becomes independent of the SCF of the specific final state, indexed by F and thereby representable as $C_{j,f}$, without any superscript. Likewise, the single-particle overlaps (see Eq. 3.15) are independent of F and can be written as ${}^F\xi_{p,q} = \xi_{p,q}$.

We dub this approximation the many-body X-ray emission spectroscopy (MBXES) method by direct analogy with a similar approach used for X-ray absorption [266–268]. The similarity between ooDFT (blue) and MBXES (green) plots in Fig. 3.1 justifies the underlying approximation for these cases.

Unlike ooDFT, the MBXES method requires only two SCF calculations, one with a core-hole (for the initial state) and one without (for all final states). It also avoids instabilities resulting from the near-degeneracy of orbitals. Additionally, unlike the response-based approaches, it can be used in plane-wave calculations in conjunction with pseudopotentials, where a modification in the pseudopotential incorporates the effects of the core-hole and thereby eliminates the need for any additional charge-constraining technique. As shown in Fig. S6, the O K α MBXES spectrum of C_6H_5OH calculated using Q-Chem, an all-electron software-package employing localized basis functions compares well with that calculated using the pseudopotential-based plane-wave software Quantum ESPRESSO [269].

It is also worth noting in Fig. 3.1 (and as we shall see later in Section 3.4.6 and in the SI) the good agreement between the de-excitation energies of ooDFT (whose energies are computed self-consistently) and MBXES (derived from the GS KS eigenvalues). For the particular functional and these systems, this implies that the KS eigenvalues are good approximations to the quasiparticle excitation energies – although there will obviously be cases, outside the examples we report, where increased accuracy can be provided by better functionals or perturbative approaches such as the GW approximation [270, 271].

To illustrate the accuracy of the more efficient MBXES approach vs. adiabatic LR-TDDFT, we provide multiple examples for the XES of small molecules in Fig. 3.2 where the corresponding change in dipole moment is also provided. Note that, consistent with the discussion presented earlier, in general, a larger change in the valence electrostatic dipole

moment upon core-ionization results in a larger difference between the two spectra, with the former typically showing better agreement with experiment. For example, for C_6H_5OH (spectra also presented in the bottom panel of Fig. 3.1), which corresponds to a relatively large change (3.01 D) in the electrostatic dipole moment, the penultimate peak at 527.5 eV in the adiabatic LR-TDDFT spectrum is seen to be the most intense. On the other hand, in the MBXES, ooDFT and experimental spectra, which show appreciable mutual agreement in the relative intensity of the peaks, the third peak from the right hand side (at 525.1 eV for the experimental spectrum and at 525.5 eV for the ooDFT and MBXES spectra) has the highest intensity. It is to be noted that, although, the simulated spectrum, especially the de-excitation energies depend, to some extent, on the exchange-correlation functional and can therefore be made to agree better with experiments by using improved functionals, the oscillator-strength accuracy, which is our primary focus here, is definitely improved over LR-TDDFT by the orbital-constrained approaches. Conversely, for CH_3Cl (spectra also presented in the top panel of Fig. 3.1), which corresponds to a much smaller change (1.43 D) of the electrostatic dipole moment, all of the spectra show an intense peak around 2817 eV and a shorter but broader feature at a lower energy of 2814 eV.

3.4.4 Single-particle approximations

3.4.4.1 Projected Ground State

Further approximations, sacrificing accuracy in lieu of computational efficiency can be made beyond MBXES. For example, we propose the projected ground-state (pGS) approximation, which takes into account the change in the participating valence orbital in response to the core hole filling, but ignores the same for the other (spectator) orbitals.

From Eq. 3.22 the MBXES transition dipole moment is given by

$$M^f_{\mathrm{MBXES}} = \sum_{j=1}^{N} C_{j,f}\tilde{o}_{0j}. \tag{3.24}$$

Now, if we approximate the determinantal overlap with the relevant single-particle counterpart,

$$\begin{aligned} C_{j,f} &= \left\langle \Psi_{-f} | \Psi_{\mathrm{I}}^{+0-j} \right\rangle \\ &= \left\langle \Psi_{\mathrm{GS}} \left| a_f^\dagger \tilde{a}_j \right| \Psi_{\mathrm{I}}^{+0} \right\rangle \\ &\approx \left\langle \phi_f | \tilde{\phi}_j \right\rangle^* \\ &= \xi^*_{j,f} \end{aligned} \tag{3.25}$$

then, the transition dipole moment becomes

$$M^f_{\mathrm{pGS}} = \sum_{j=1}^{N} \xi^*_{j,f}\tilde{o}_{0j} \tag{3.26}$$

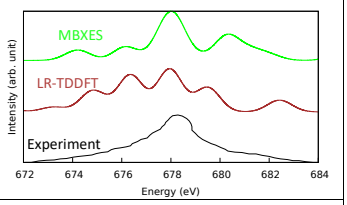
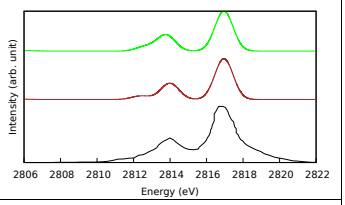
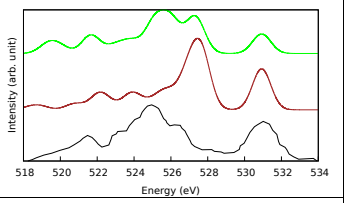
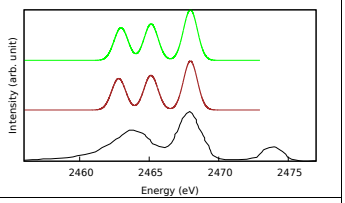
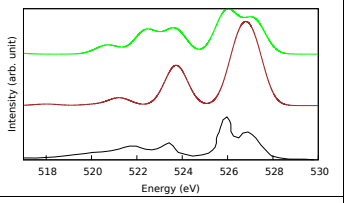
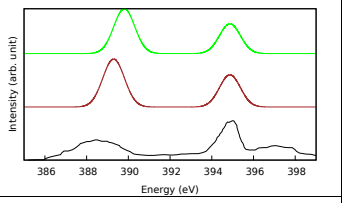
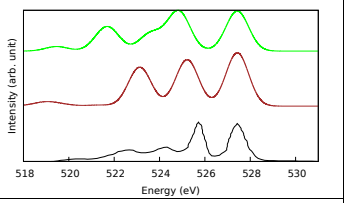
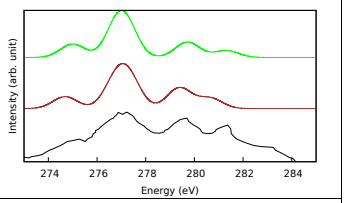
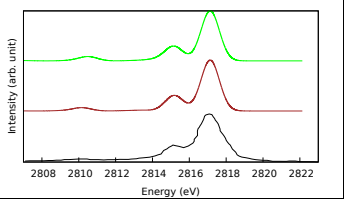
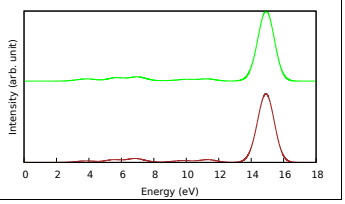
Molecule	ΔD	Spectra	Molecule	ΔD	Spectra
C_6H_5F [F K α]	3.28		CH_3Cl [Cl K β]	1.43	
C_6H_5OH [O K α]	3.01		H_2S [S K β]	0.74	
$(CH_3)_2SO$ [O K α]	2.72		NH_3 [N K α]	0.45	
$(CH_3)_2CO$ [O K α]	2.57		CH_3OH [O K α]	0.3	
CF_3Cl [Cl K β]	1.75		C_6H_5OH [Valence De-excitation from initial state with hole in 11th KS orbital]	0.17	

Figure 3.2: Table showing experimental [261, 262, 272–275] (black) adiabatic LR-TDDFT (brown) and MBXES (green) spectra of different molecules. The second column shows the change in the electrostatic dipole moment, in Debye units, of the remaining electrons following the relevant core-ionization.

while $\left|\phi_f^{\mathrm{pGS}}\right\rangle = \sum_{j=1}^{N} \xi_{jf}^{*} \left|\tilde{\phi}_j\right\rangle$ gives the corresponding auxiliary orbital.

Significant computational savings arise here due to replacing the N-electron determinant with a single overlap matrix element. In terms of the physical approximation involved, pGS takes into account the fact that the participating orbital (i.e., the valence orbital from which the electronic de-excitation takes place) should be different in presence and in absence of the core-hole. What is missing in the pGS approximation is the explicit inclusion of relaxation of the other valence electrons. In general, absorption spectra calculated with the pGS approximation (see Supplemental Material for Ref. 265 for simulated examples), show good agreement with MBXES spectra, despite this simplification.

3.4.4.2 Ground State

Note that in all of the aforementioned approximations, separate sets of orbitals $\{\tilde{\phi}_j\}$ and $\{\phi_j\}$ (or $\{\phi_j^{(f)}\}$) are used for the initial and the final state, indicating separate SCFs for states with and without a core-hole. As a final approximation within the orbital-constrained framework, we now neglect this effect by extending the sum in Eq. 3.26 to all orbitals, thereby approximating the transition moment entirely within a *Ground State Single Particle* (GS) treatment as follows:

$$\begin{aligned}\left(M_{\mathrm{GS}}^{f}\right)^{*} &= \sum_{j=1}^{\mathrm{all}} \xi_{jf} \left\langle \tilde{\phi}_j \left|\hat{o}\right| \tilde{\phi}_0 \right\rangle \\ &= \langle \phi_f | \sum_{j=1}^{\mathrm{all}} \left|\tilde{\phi}_j\right\rangle \left\langle \tilde{\phi}_j \left|\hat{o}\right| \tilde{\phi}_0 \right\rangle \\ &= \left\langle \phi_f \left|\hat{o}\right| \tilde{\phi}_0 \right\rangle \approx \langle \phi_f \left|\hat{o}\right| \phi_0 \rangle\end{aligned} \tag{3.27}$$

where the approximation that the core orbital remains unchanged, $|\phi_0\rangle \approx \left|\tilde{\phi}_0\right\rangle$, is validated by our calculations showing $\left\langle \phi_0 | \tilde{\phi}_0 \right\rangle \approx 1$. A table summarizing the different approximations within the framework of orbital-constrained XES is presented in Tab. 3.1.

3.4.4.3 Relation between $M_{\mathbf{MBXES}}^{f}$ and $M_{\mathbf{GS}}^{f}$

It is instructive to investigate analytically the conditions for equivalence of the GS and the MBXES spectra. To this end, we aim to find the linear-algebraic relation between the MBXES coefficients ($\{C_{i,j}\}$) and the single-particle overlaps($\{\xi_{i,j}\}$). The inverse of the lowest $N \times N$ overlap matrix can be related to its matrix of cofactors as

$$\xi^{-1} = \frac{\mathrm{adj}(\xi)}{\det(\xi)} = \frac{\mathbf{C}^{T}}{\det(\xi)}. \tag{3.28}$$

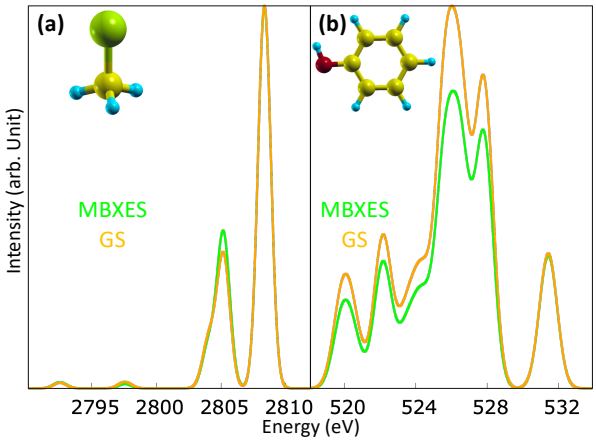


Figure 3.3: Emission spectra calculated using the MBXES (green line) and the GS (orange line) approach. Panel (a) shows Cl $K\beta$ emission from CH_3Cl, which is associated with a 1.43 Debye change in valence dipole moment, while panel (b) shows O $K\alpha$ emission from C_6H_5OH, for which the valence dipole-moment changes by 3.01 Debye. Yellow, cyan, red and green spheres denote C, H, O and Cl atoms, respectively. Each GS spectrum has been multiplied by a constant factor so that the intensity corresponding to highest-energy peak matches the MBXES counterpart.

So, for each matrix element,

$$\left(\xi^{-1}\right)_{f,j} = \frac{C_{j,f}}{\det(\xi)} \tag{3.29}$$

Therefore, if the $(N \times N)$ matrix ξ is orthogonal (i.e., if its transpose is equal to its inverse), then, noting that the determinant of an orthogonal matrix is ± 1, we obtain

$$\xi^*_{j,f} = \pm C_{j,f}, \tag{3.30}$$

and so, ignoring the ambiguity in sign,

$$(M^f_{\mathrm{MBXES}})^* \approx \sum_{j=1}^{N} \xi_{j,f} \left\langle \tilde{\phi}_j \left| \hat{o} \right| \tilde{\phi}_0 \right\rangle, \tag{3.31}$$

which is the projected ground-state approximation. And, furthermore, if the overlap matrix ξ is truly orthogonal, then this implies that the initial and final state occupied subspaces overlap perfectly and $\xi_{j,f} = 0$ for $j > N; f \leq N$. So, we can extend the sum above over all j and obtain

$$|M^f_{\mathrm{MBXES}}|^2 = |M^f_{\mathrm{GS}}|^2 \tag{3.32}$$

To reiterate, this is a remarkable simplification, which can be restated as follows: If the electrons occupy the same space in the ground and the core-excited state, then the GS and the MBXES spectra will be identical. Thus, the GS approach is likely to be inaccurate if the core-ionization induces a large change in the valence electron density. Changes in electronic density in response to perturbations are proportional to the electronic polarizability, and so the associated polarization response seems like an appropriate metric for assessing this equivalence between GS and MBXES spectral intensity.

This is reflected in Fig. 3.3, which displays the GS and the MBXES spectra for CH_3Cl Chlorine $K\beta$ and C_6H_5OH Oxygen $K\alpha$ emissions. In order to facilitate comparison, for each system the GS spectrum is scaled by a constant factor to match the intensity of the last peak, which essentially amounts to equating the oscillator strengths of the highest energy transition. Fig. 3.3 reveals that for CH_3Cl the two spectra are nearly-coincident, which is not the case for C_6H_5OH, indicating, for the latter, a disagreement among the relative oscillator strengths of the different transitions computed within the GS and the MBXES approximation. In accordance with the aforementioned analytical argument, this disagreement can be attributed to the larger change in valence electronic polarization in C_6H_5OH.

For context, the most common practice of using valence-projected density of states (pDOS) to interpret XES [276–279] is yet a further approximation that ignores the single-particle matrix elements.

We emphasize here that, even though the inaccuracies of the adiabatic LR-TDDFT and the GS formalism can be associated with the same metric, they, in fact, stem from quite different effects. In the former, the drawback lies in a partial description of the *final* (core-filled) state while in the latter it can be attributed to a complete neglect of the *initial* (core-ionized) state effects. However, in both cases, the shortcoming lies in the inability to adequately account for the difference between the valence electronic structure of the initial and the final state. Therefore, the aforementioned metric, which provides a measure for this difference is applicable for both of these diverse approaches.

3.4.5 Auxilliary orbitals

The auxiliary orbitals can provide a visual justification for the similarity/difference in the spectra seen in Fig. 3.1. As a representative example, in Fig. 3.4 we show the auxiliary orbitals corresponding to the 11^{th} de-excitation in CH_3Cl (Cl $K\beta$) and the 23^{rd} de-excitation in C_6H_5OH (O $K\alpha$)(see also Fig. S1). While for CH_3Cl, all three orbitals are in good agreement, for C_6H_5OH noticeable differences can be spotted in the auxiliary orbital obtained using LR-TDDFT. Note, however, that all three auxiliary orbitals on the right-hand-side panels are qualitatively similar, indicating that they contain appreciable contribution from the same set of constituent initial state orbitals (see also Fig. S5 in the SI which shows a bar-plot of the coefficients of the contributing initial-state orbitals). However, as a consequence of the initial state bias, $\phi_{23}^{\mathrm{TDDFT}}$ is skewed toward the core hole on the oxygen atom, while the ooDFT and MBXES counterparts display more delocalized nature. As expected, this leads to a difference in the associated oscillator strengths.

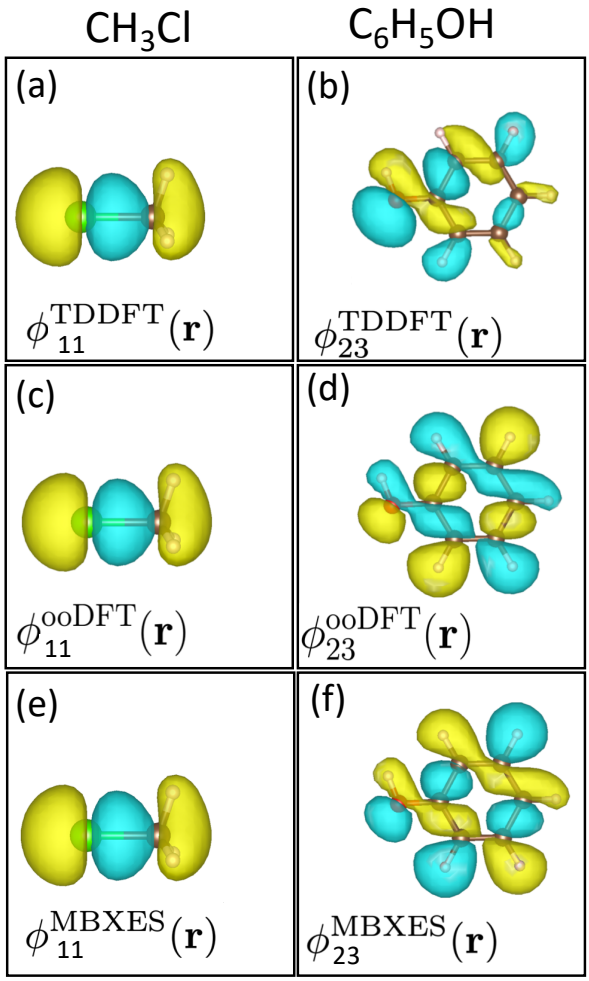


Figure 3.4: Panel (a), (c) and (e) display the auxiliary orbital, evaluated using respectively LR-TDDFT, ooDFT and MBXES associated with the 11-th de-excitation in the Cl K β emission of CH_3Cl. Panels (b) , (d) and (f) plot the same for the 23-rd de-excitation in the O K α emission of C_6H_5OH. These plots demonstrate that MBXES effectively approximates ooDFT while, for phenol, LR-TDDFT exhibits errors associated with the adiabatic approximation.

Method	Final orbitals consistent with valence hole	Relaxation of spectator electronic orbitals	Initial orbitals consistent with core-hole	Auxiliary orbital
ooDFT	Yes	Yes	Yes	$\sum_{j=1}^{N} C_{j,f}^{(f)} \left\vert \tilde{\phi}_j \right\rangle$
MBXES	No	Yes	Yes	$\sum_{j=1}^{N} C_{j,f} \left\vert \tilde{\phi}_j \right\rangle$
pGS	No	No	Yes	$\sum_{j=1}^{N} \xi_{jf}^{*} \left\vert \tilde{\phi}_j \right\rangle$
GS	No	No	No	$\sum_{j=1}^{\text{all}} \xi_{jf}^{*} \left\vert \tilde{\phi}_j \right\rangle$

Table 3.1: Table summarizing important properties and approximations corresponding to various simulation-methods discussed in the paper. Note that, in each case, the transition-dipole moment is given by $\langle \text{core} \left\vert \hat{o} \right\vert \text{Auxiliary Orbital} \rangle$, where $\hat{o}$ is the single-particle dipole operator.

3.4.6 Extension to valence de-excitations

By contrast with XES, for the majority of low energy valence electronic transitions, the response-based methods have proven to be highly successful for decades [280]. In reference to the metric introduced above, it is worth noting that, in most cases, such transitions are usually associated with a comparatively small change in the polarization owing to the relatively delocalized nature of the valence hole (compared to the atomic core hole). As an illustrative example, in Fig. 3.5, we plot the simulated emission spectra of C_6H_5OH associated with de-excitation from an initial state with a hole in the 11-th (valence) KS orbital. Owing to the small (0.17 D) change in the electrostatic dipole moment, the LR-TDDFT result is found to be in excellent agreement with the ooDFT and the MBXES counterparts. In contrast, the charge-transfer excitation, which is accompanied by an appreciable spatial separation of charges resulting typically in a large change in the valence polarization, is a well-documented example of the failure of adiabatic LR-TDDFT in predicting the excitation-energy [281, 282], although further studies are needed to investigate the accuracy in the calculation of oscillator strength.

3.5 Conclusions

In summary, with a focus on valence-to-core electronic de-excitation, we use a generalized formalism to link the disparate response-based and orbital-constrained approaches, which are the two standard pathways for modeling electronic excited states starting from DFT

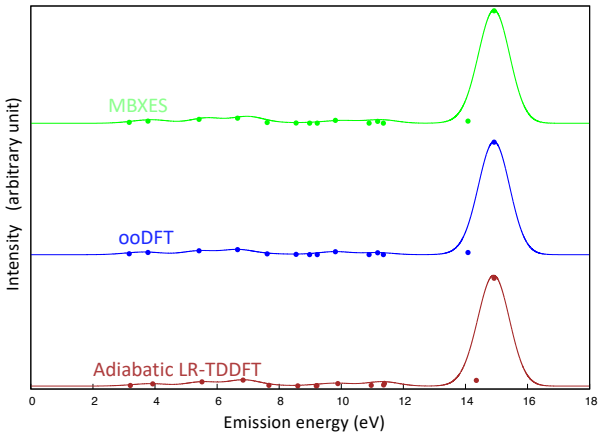


Figure 3.5: Simulated emission spectra, computed with MBXES, ooDFT and adiabatic LR-TDDFT, corresponding to electronic de-excitation of C_6H_5OH from an initial state with a hole in the 11-th KS valence orbital.The associated change in valence polarization is 0.17 D.

calculations. Our analytical and computational results show that, unless the de-excitation induces a small change in the valence polarization, the popular adiabatic approximation of the response-based approach might incur significant inaccuracies in the oscillator strength. Consequently, such a change in polarization can be used as an effective metric for assessing the validity of the adiabatic approximation. In fact, a valence-to-core de-excitation can often render the adiabatic approximation inadequate since the filling of a localized core orbital may lead to appreciable rearrangement of the valence electrons. The aforementioned shortcoming is not present in the orbital-constrained approach such as ooDFT, which typically yields accurate results. By modifying the ooDFT formalism with physically-motivated approximations, we propose a flexible computational technique, MBXES, which can simulate oscillator strengths accurately at a much lower computational cost.

Supplementary Material

Supplementary information for this work can be accessed via Ref. 265. They include additional plots for some of the simulated systems.

Chapter 4

Relativistic Orbital Optimized Density Functional Theory for Accurate Core-Level Spectroscopy

4.1 Introduction

Spectroscopy of core-level electrons with X-rays is a convenient and popular tool for studying chemical systems. A specific core-level of a given element normally has a characteristic energy that is quite distinct from inner-shells of other elements, making the technique element specific. Furthermore, core electrons do not play a direct role in chemical bonding, and are thus effectively localized around the nucleus. Spectroscopic probe of these electrons therefore yields information about local chemical environment of individual atoms. Core-level spectra can thus yield useful information about the local coordination environment [283, 284], extent of covalency in ligand-metal interactions [285, 286] or the oxidation state [287]. Time-resolved core-level spectroscopy can also be used as a probe to study photoinduced chemical dynamics [288–290].

Computational simulations of core-level spectra are useful for gaining insight into experiment, and potentially identifying new species whose signature may appear in transient spectra [290–292]. Traditional quantum chemistry methods for excited states [89, 91] are however quite challenged by this task, especially since such techniques are mostly developed for (and validated on) problems involving only valence electrons. For example, the widely used linear-response time-dependent density functional theory (TDDFT) approach [91, 293] cannot adequately describe the relaxation of the core hole. This leads to rather large errors of $\sim$ 10-20 eV for the K-edge (1s orbitals) of C, N, O and F [292, 294–298], if highly specialized functionals [299, 300] are not employed. Heavier elements lead to even larger errors, such as $\sim$ 50 eV for the P, S, Cl K-edges [301–303] and $>$ 100 eV for the Fe K-edge [304]. TDDFT spectra therefore usually need to be empirically translated by many eVs, in order to align with experiment [292, 302, 304, 305]. Similar behavior is observed for the equation-of-motion

coupled cluster singles and doubles (EOM-CCSD) method [89, 306], although the shift required is typically much smaller (< 2 eV for second period elements) [307–311]. EOM-CCSD is however quite computationally demanding, with the computational cost scaling as $O(N^6)$ vs system size N (compared to $O(N^{3-4})$ for DFT).

Orbital-optimized (OO) methods optimize orbitals for each excited state individually, and separately from those of the ground state. OO can therefore effectively model the relaxation of the core hole, leading to much better agreement with experiment [238, 312–314] without any need for empirical shifts. Unfortunately, OO methods had been historically underutilized due to a risk of 'variational collapse', in which the calculation converges to a lower energy state (often the ground state) instead of the desired high energy excitation. However, there has been considerable recent interest in excited state OO, resulting in many new approaches that aim to reliably converge to any chosen state without the risk of variational collapse [243, 315–320]. In practice, OO-DFT methods require more compute time than TDDFT if a large number of states are desired, such as in (near-)degenerate bands. This stems from OO-DFT having to iteratively optimize multiple states individually while TDDFT can simultaneously compute them. However, OO-DFT retains the same computational scaling as ground state DFT or TDDFT. An overview of the successes and challenges with OO-DFT methods can be found in Ref 238. OO-DFT methods are thus increasingly being employed to study core-level spectra [319, 321–325], with the modern SCAN [72] functional leading to very low error [321, 322, 325, 326] (< 1 eV) vs experiment for the K-edge of C, N, O and F, as well as L-edge of Si, P, S and Cl.

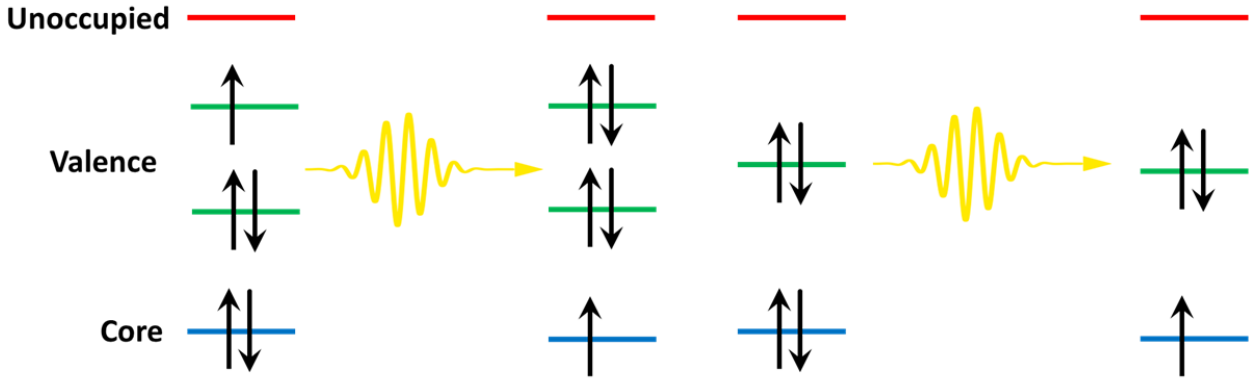
The K-edge of elements heavier than F however cannot be as accurately modeled with non-relativistic quantum mechanics. Naive use of the Bohr atom model suggests that the speed of 1s electrons would scale linearly with the atomic number Z, eventually attaining the speed of light at $Z > \alpha^{-1} \approx 137$ (where α is the fine structure constant). Relativistic effects become perceptible at much smaller Z, with calculations indicating that non-relativistic quantum mechanics underbinds the 1s electrons of Ne by 1 eV [327]. It is therefore necessary to incorporate relativistic effects into OO-DFT, if < 1 eV error vs experiment is desired for computed K-edge spectra of third period elements and beyond. Scalar relativistic treatment is however often overlooked for linear-response TDDFT, as the ad-hoc empirical shifts (typically larger than the relativistic correction) utilized to align computation with experiment account for it to some extent [226, 328, 329]. Nonetheless, explicit use of relativistic effects in TDDFT has been previously explored, albeit mostly within a real-time framework [226, 295, 330]. Similarly, both empirical shifts [308] and explicit inclusion [121, 331] have been used to account for relativistic effects in coupled cluster methods. The use of element-specific corrections to nonrelativistic TDDFT results has also been examined in the past [299].

In this work, we utilize the spin-free exact-two component one electron (SFX2C-1e, henceforth referred to as just X2C) model for relativistic quantum chemistry [127, 128, 332–338] to obtain improved OO-DFT core-level spectra. The results indicate that OO-DFT/X2C is quite accurate for third period elements, permitting semiquantitative prediction of core-level spectra at ground state DFT cost.

4.2 Brief overview of theory

The X2C model transforms the one particle terms of the electronic Hamiltonian (i.e. kinetic energy and external potential) via the solutions of the four component, one electron Dirac Hamiltonian. The two-particle (i.e. interelectron interaction) terms are treated within the pure Coulomb formalism and, therefore, are left unaltered in the non-relativistic form, permitting straightforward application of DFT. The transformation is briefly described in the supporting information, and we invite interested readers to examine Refs. 128 and 336 for further details regarding X2C.

Using the X2C transformed one particle Hamiltonian, we obtained the ground state energy via the standard Kohn-Sham [58] (KS) formalism. Excited states are more challenging, as many excitations unpair electrons and therefore require multiple Slater determinants for a spin-pure description. DFT for such states is not straightforward, as the KS formalism is single determinant by construction. We therefore utilize three related OO-DFT ansatze for modeling three classes of excited states. These ansatze are described in detail in Ref 238, but we provide a brief outline here for convenience.



(a) Core→SOMO excitation in a radical. (b) Core-ionization of closed-shell system

Figure 4.1: Schematic for processes where ΔSCF is appropriate.

A state with no unpaired electrons, or one with all unpaired electrons of the same spin can be represented by a single Slater determinant. It is straightforward to optimize a single Slater determinant with an excited state electronic configuration (apart from the aforementioned risk of variational collapse). This protocol is called ΔSCF [339, 340] and is suitable for states that result from 1s→ SOMO (singly occupied molecular orbital) transitions of open-shell species, or core-ionized states of closed-shell molecules (as shown in Fig 4.1). Relativistic Δ Hartree-Fock (HF) has indeed been used to study core-ionization energies [341, 342]. However, ΔSCF is not appropriate for singly excited singlet excited states of closed-shell molecules, as both the up and down spins are equally likely to be excited (as shown in Fig 4.2). Exciting only one spin results in a spin-contaminated determinant midway between

singlet and triplet. Spin-contaminated ΔSCF energies have nonetheless been utilized in the past for core-excitation energy calculations, with element-specific relativistic corrections for heavy elements [312].

Restricted open-shell Kohn-Sham [343, 344] (ROKS) obtains a pure singlet energy by spin-projection on the spin-contaminated determinant. ROKS is consequently the optimal OO-DFT approach for singlet excited states with two unpaired spins, although it cannot be applied if there are more than two unpaired spins. Such states require a more general recoupling scheme described in Refs 238 and 322. This is however only necessary for transitions from the core to completely unoccupied levels in open-shell systems, with ΔSCF and ROKS being sufficient for all the states considered in this work. OO-DFT therefore encompasses ΔSCF and methods like ROKS and the general recoupling scheme that derive from it.

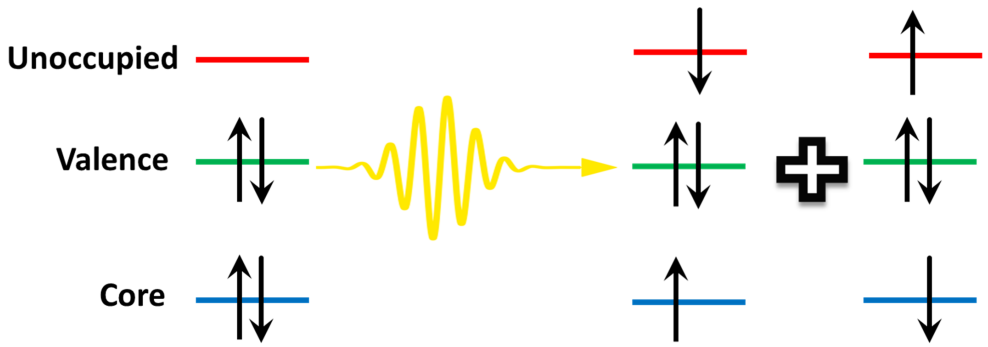


Figure 4.2: Schematic for a singlet core excitation in a closed-shell species. Two open-shell determinants (as seen on the right) are equally likely, and are individually halfway between singlet and triplet in character. ROKS is thus essential for spin-purity of the excited state.

4.3 Computational methods

All calculations were performed with a development version of the Q-Chem 5.4 package [263] and these new capabilities will be publicly available with the next release of code. Local exchange–correlation integrals for DFT were calculated over a radial grid with 99 points and an angular Lebedev grid with 590 points. Ref 322 lays out the protocol for computing ΔSCF excited states, while Ref 321 does the same for ROKS. The restricted open-shell optimizations necessary in this process were performed through square gradient minimization (SGM [318]) and unrestricted optimizations with initial maximum overlap method (IMOM [315]). The core-hole was localized onto a single atom for species with equivalent atoms (like S in CS_2), in order to prevent errors arising from delocalization [74, 345] of the hole over multiple sites [321]. Standard values of dielectric constant ϵ_r and refractive index n were used for IEF-PCM modeling of common solvents like cyclohexane (listed in the supporting information).

The corresponding data for most solid state materials was not available, and consequently all solid state environments were modeled with NaCl parameters ($\epsilon_r = 6, n = 1.5$). This should be reasonable for ionic solids, although perhaps a little too polar for molecular solids like $OPPh_3$ (Table 4.3). However, the resulting values for $OPPh_3$ were quite close to vacuum calculations (as shown in the supporting information), suggesting that the environment exerted negligible impact on the spectrum.

Experimental geometries were used whenever possible, through gas phase data from NIST [346] or crystal structures from the Cambridge structural database [347]. Structures were optimized with ωB97M-V [79]/aug-pcseg-1 under gas phase conditions, if experimental data was unavailable. The source of all of the geometries is listed in the supporting information, along with the associated atomic coordinates. In particular, the $CuCl_4{}^{2-}$ ion studied is of D_{2h} (distorted tetrahedral) symmetry, corresponding to Cs_2CuCl_4 [348]. The ground state geometries were employed for excited state calculations, consistent with the Franck-Condon principle [349, 350].

4.4 Results and Discussion

We first examined the performance of OO-DFT/X2C in predicting the gas phase K-edge spectra of the third period elements (and Ne) with Table 4.1 reporting ΔSCF 1s electron binding energies for several closed-shell species. All presented functionals have root mean square error (RMSE) $<$ 1 eV vs experimental X-ray photoelectron spectra (XPS). These functionals were identified via screening across many functionals over a smaller set of species (SiH_4,PH_3,H_2S,HCl and Ar). This screening also revealed that other well known functionals like B3LYP [77, 363], PBE0 [78], or TPSS [71] have larger errors ($\sim$ 1-3 eV, as shown in the supporting information) that nonetheless represent a major improvement over TDDFT or non-relativistic OO-DFT. Out of the selected functionals, SCAN fares particularly well, yielding an RMSE of 0.4 eV and a maximum deviation of 0.8 eV from experiment. X2C is crucial for this level of agreement, as SCAN with the non-relativistic (NR) Hamiltonian leads to errors of several eV (as shown in Table 4.1). The related SCANh [364] functional performs slightly worse, but is still fairly accurate. SCANh does have positive mean error (ME), indicating it systematically overestimates the binding energy. This overestimation is a consequence of the presence of HF exchange (10%) in the functional, as pure HF overestimates by $\sim$ 2 eV. Overestimation is more evident for SCAN0 [365] (which has 25% HF exchange) and BHHLYP [366] (50% HF exchange). However, it is important to note that the ME is strongly influenced by the choice of the local exchange-correlation model. For example, functionals based on PBE [67] appear to be far more sensitive to % HF exchange, than ones derived from SCAN (as seen in the supporting information). In addition, most local functionals strongly underbind core-electrons, and would require admixture of a very large amount of % HF exchange to have low error (BHHLYP being a prominent example). SCAN is a notable exception in this regard, as it has low error despite being a local functional.

We next considered prediction of X-ray absorption spectra (XAS) with ROKS, which is

Molecule	Expt.	SCAN (NR)	SCAN	SCANh	SCAN0	BHHLYP	HF
Ne	870.2 [351]	869.3	870.3	870.3	870.3	870.5	869.6
Mg	1311.5 [352]	1309.3	1311.6	1311.6	1311.6	1311.5	1311.4
SiH_4	1847.0 [353]	1842.7	1847.1	1847.2	1847.3	1847.3	1848.0
SiF_4	1852.5 [353]	1847.8	1852.1	1852.4	1852.7	1852.8	1853.7
$\mathbf{Si}Cl_4$	1850.6 [353]	1846.1	1850.5	1850.7	1850.9	1851.0	1852.0
$SiBr_4$	1849.7 [353]	1845.5	1849.8	1850.0	1850.3	1850.3	1851.3
PH_3	2150.9 [354]	2145.1	2151.0	2151.1	2151.2	2151.0	2152.0
PF_3	2156.4 [354]	2150.0	2155.8	2156.1	2156.4	2156.4	2157.7
PF_5	2159.4 [354]	2153.2	2159.0	2159.3	2159.8	2160.0	2161.6
POF_3	2157.8 [354]	2151.7	2157.5	2157.8	2158.2	2158.4	2159.9
H_2S	2478.5 [355]	2470.7	2478.4	2478.6	2478.8	2478.4	2479.5
CS_2	2478.1 [356]	2470.4	2478.2	2478.3	2478.5	2478.0	2479.1
SF_4	2486.9 [357]	2478.7	2486.5	2486.8	2487.3	2487.4	2489.3
SF_6	2490.1 [355]	2481.9	2489.6	2489.9	2490.5	2490.6	2492.7
SO_2	2483.7 [355]	2475.9	2483.6	2483.9	2484.3	2484.2	2486.0
CSO	2478.7 [356]	2471.2	2479.0	2479.1	2479.3	2478.8	2479.9
$\mathbf{S}F_5Cl$	2488.9 [357]	2480.9	2488.6	2489.0	2489.5	2489.5	2491.5
HCl	2829.8 [358]	2820.3	2830.3	2830.4	2830.6	2830.0	2831.4
Cl_2	2830.2 [358]	2820.8	2830.8	2831.0	2831.2	2830.5	2831.9
CH_3Cl	2828.4 [261]	2819.2	2829.2	2829.3	2829.5	2828.9	2830.2
$SF_5\mathbf{Cl}$	2829.6 [359]	2820.3	2830.4	2830.6	2830.9	2830.3	2831.8
CCl_3F	2829.3 [261]	2820.0	2830.0	2830.2	2830.4	2829.8	2831.2
Ar	3206.3 [360]	3194.1	3206.9	3207.0	3207.3	3206.5	3208.1
RMSE		7.4	0.4	0.5	0.6	0.4	1.7
ME		-6.9	0.1	0.3	0.5	0.3	1.5
MAX		12.2	0.8	1.0	1.3	0.7	2.7

Table 4.1: Gas phase XPS K-edge binding energies for Ne and third period elements (in eV). Computed values were found from restricted open-shell ΔSCF calculations, using the aug-pcX-2 basis [361] when available and decontracted aug-pcseg-2 [362] for H/Br. Non-relativistic (NR) values from SCAN are also provided for comparison. The root mean square error (RMSE), mean error (ME) and maximum absolute error (MAX) are also reported. The atomic site of the ionization is **bolded** when multiple possibilities exist.

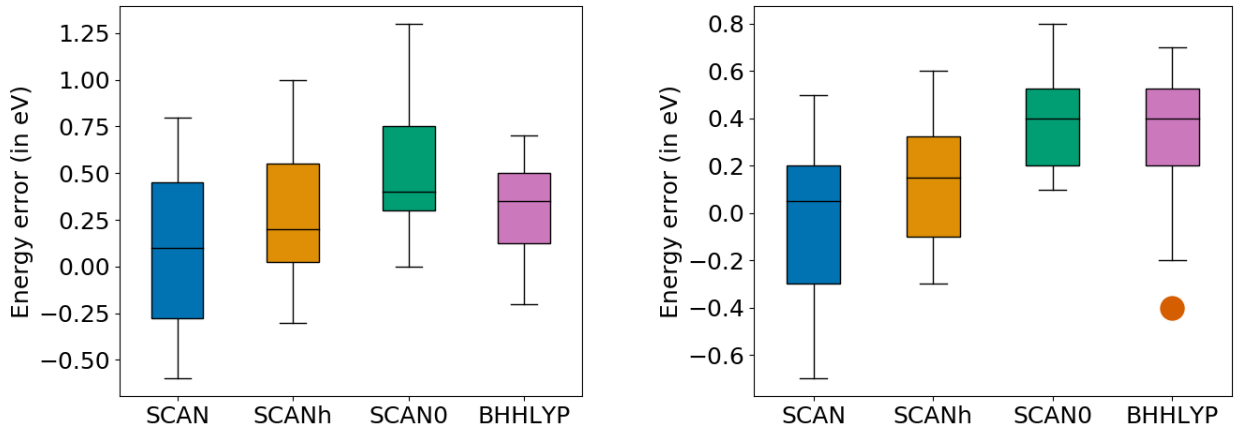
quite effective in predicting singlet core excitation energies of second period elements [321]. Table 4.2 shows that inclusion of scalar relativistic effects through the X2C model permits high accuracy for third period elements as well. SCANh yields the best performance with an RMSE of 0.3 eV and a maximum absolute error (MAX) of only 0.6 eV. SCAN and SCAN0 also yield quite good performance. In fact, the RMSE for all the presented functionals

Molecule	Expt.	SCAN (NR)	SCAN	SCANh	SCAN0	BHHLYP	HF
SiH_4	1842.7 [353]	1838.56	1842.9	1843.1	1843.3	1843.4	1844.9
SiF_4	1849.0 [353]	1844.13	1848.5	1848.8	1849.1	1849.4	1851.3
$\mathbf{Si}Cl_4$	1846.0 [353]	1841.37	1845.7	1845.9	1846.2	1846.5	1848.5
$SiBr_4$	1845.0 [353]	1840.58	1844.9	1845.1	1845.4	1845.6	1847.7
PH_3	2145.8 [367]	2140.15	2146.0	2146.2	2146.4	2146.3	2148.1
PF_3	2149.3 [367]	2143.55	2149.4	2149.6	2149.9	2150.0	2151.9
PF_5	2155.0 [367]	2148.59	2154.5	2154.7	2155.2	2155.5	2159.8
POF_3	2153.3 [367]	2147.16	2153.0	2153.3	2153.7	2153.9	2158.0
H_2S	2472.7 [368]	2465.05	2472.8	2473.0	2473.2	2472.9	2475.0
CS_2	2470.8 [356]	2463.53	2471.3	2471.3	2471.4	2471.0	2472.5
SF_4	2477.3 [369]	2469.71	2477.4	2477.7	2478.0	2478.0	2480.5
SF_6	2486.0 [368]	2477.55	2485.3	2485.7	2486.2	2486.5	2490.1
SO_2	2473.2 [368]	2465.53	2473.3	2473.4	2473.6	2473.4	2475.5
CSO	2472.0 [356]	2464.73	2472.5	2472.6	2472.7	2472.2	2473.9
$\mathbf{S}F_5Cl$	2483.5 [359]	2475.11	2482.8	2483.2	2483.7	2483.9	2487.3
HCl	2823.9 [358]	2813.79	2823.8	2824.0	2824.2	2823.7	2825.7
Cl_2	2821.3 [358]	2811.05	2821.1	2821.2	2821.4	2820.9	2822.8
CH_3Cl	2823.5 [261]	2813.60	2823.6	2823.8	2824.0	2823.6	2825.8
$SF_5\mathbf{Cl}$	2821.8 [359]	2811.83	2821.8	2821.9	2822.1	2821.6	2823.7
CCl_3F	2822.8 [261]	2813.19	2823.2	2823.4	2823.6	2823.2	2825.3
RMSE		7.6	0.4	0.3	0.5	0.5	2.8
ME		-7.3	-0.1	0.1	0.4	0.3	2.7
MAX		10.3	0.7	0.6	0.8	0.7	4.8

Table 4.2: Lowest dipole allowed gas-phase XAS excitation energy for third period elements (in eV). Computed values were found from ROKS, using the aug-pcX-2 basis when available and decontracted aug-pcseg-2 for H/Br. NR values from SCAN are also provided for comparison. The atomic site of the ionization is **bolded** when multiple possibilities exist.

is comparable to the typical experimental energy resolution of $\sim$ 0.5 eV, and therefore indicative of semi-quantitative performance. Curiously, SCAN significantly underestimates the excitation energy for highly fluorinated compounds (SF_6, CF_3SF_5 etc.), highlighting a potential limitation for this otherwise excellent performing local functional. This systematic underestimation is partially mitigated with HF exchange, leading to SCANh performing somewhat better. On the other hand, SCAN0 has a systematic bias towards overestimation due to a greater part of HF exchange being present. SCANh therefore offers a reasonable middle path, although it would perform poorly for cases where SCAN already overestimates or SCAN0 underestimates. We also note that our RMSEs are considerably smaller than the several eV errors reported by an earlier study [338] using relativistic orthogonality constrained

DFT (OC-DFT) [242], which may in part stem from use of B3LYP in that work.



(a) K-edge ionization.

(b) Lowest dipole allowed K-edge excitation.

Figure 4.3: Box plot for errors (vs experiment) in computed values reported in Tables 4.1-4.2.

Fig 4.3 visually summarizes the key result shown in Tables 4.1 and 4.2, namely that OO-DFT/X2C is effective in predicting core-level excitation/ionization energies of third period atoms in isolated small molecules. The error distributions are quite compact overall, with the typical range being below 1 eV. In particular, the SCAN, SCANh and BHHLYP functionals typically have errors below 0.5 eV, and never over 1.0 eV (for the species considered). They therefore appear to be promising routes for prediction of gas-phase core-level spectra.

It is also worthwhile to consider larger systems in order to gauge feasibility of OO-DFT/X2C for widespread practical use. However, certain computational challenges need to be considered along the way. Decontracted basis sets of at least triple-ζ quality (ideally of the Jensen pcX-n [361] or Dunning cc-pCVnZ [370] type) appear to be necessary for computation of core-level spectra, both to account for relaxation of the core hole and for convergence of relativistic effects. However, using such bases for all atoms would be quite computationally demanding. The local nature of the core-excitation permits use of a mixed basis strategy [321] in which the decontracted triple-ζ basis is only used for the atom whose core-electrons are being probed, while the corresponding contracted double-ζ basis is sufficient for all other atoms. We have verified that this mixed basis strategy does not lead to any significant change in RMSE of SCAN/SCANh for the species in Tables 4.1 and 4.2 (shown in the supporting information) and have employed this strategy for the calculations reported hereon.

In addition, XAS for larger systems is often collected in the solid state or in solution, making it necessary to model the effect of the environment on the spectrum. The locality of core-electrons suggests that only the first coordination shell needs to be considered atomistically for the first few ('pre edge') peaks, with a continuum dielectric model being adequate for the remainder of the environment. Such polarizable continuum models (PCM [371]) are likely to be effective for species in which the core electron is excited to a valence level, but

would probably be insufficient for Rydberg-like excitations without atomistic modeling of a rather large region around the core-hole. We have employed the integral equation formalism (IEF-) PCM [372] model to account for environment effects, and thereby investigate how spectra are affected by the phase of the system [373]. IEF-PCM is likely adequate if it only induces a small shift relative to vacuum results, but more sophisticated embedding techniques might prove necessary if there is a significant difference between IEF-PCM and vacuum calculations.

Molecule	Environment	Expt.	SCAN	SCANh
Si$(Me)_4$	Gas	1843.6 [374]	1843.5	1843.7
Si$(OMe)_4$	Gas	1845.9 [374]	1845.9	1846.1
$(CH_3O)_2$**P**(S)Cl	Gas	2150.2 [367]	2150.0	2150.2
$\mathbf{P}_4O_6$	Gas	2147.5 [375]	2147.8	2148.0
O**P**Ph_3	Solid	2147.3 [301]	2147.4	2147.7
$CF_3\mathbf{S}F_5$	Gas	2483.8 [376]	2483.2	2483.6
$CH_3\mathbf{S}SCH_3$	Gas	2471.6 [377]	2472.0	2472.1
(4-Me)C_6H_4**S**H	Cyclohexane	2472.5 [291]	2472.5	2472.7
(4-Me)C_6H_4**S**·	Cyclohexane	2467.0 [291]	2467.4	2467.5
Ti**Cl**$_4$	Toluene	2821.6 [378]	2821.3	2821.2

Table 4.3: Lowest dipole allowed XAS excitation energy for slightly larger species (in eV) from experiment and theory. A mixed basis (aug-pcX-2 on excitation site, aug-pcseg-1 on all other atoms) was utilized for the calculations.

Table 4.3 shows performance for the mixed basis protocol for molecular systems in the gas phase, or non-polar solvents like cyclohexane (modeled with IEF-PCM, if present). ROKS was used for all closed-shell systems, while single determinant spin-unrestricted ΔSCF was sufficient [322] for the S 1s $\rightarrow$ SOMO transition of the 4-methylthiophenoxy ((4-Me)C_6H_4**S**) radical. Our approach appears to be quite accurate in predicting experimental energies, indicating that the OO-DFT/X2C approach can be applied to large molecules for prediction of heavy element K-edges. We also revisited earlier work on light elements [322], and demonstrated that inclusion of X2C does not cause any degradation of performance in predicting excitation energies (as shown in the supporting information).

Although we have only considered neutral species so far, core-level spectra of ionic moieties have also been collected in many experiments. These species offer an interesting regime for both testing the efficacy of our approach and for gauging environment effects in general. Table 4.4 presents a comparison between experiment and theory for Cl K-edges of several ionic species. The ions in Table 4.4 can be broadly categorized into two categories. The first are closed-shell species where Cl has a formally positive oxidation state (ClO_4^- etc) and the lowest excitation is 1s $\rightarrow \sigma^*_{\text{O-Cl}}$. These σ^* orbitals are more 'Cl like' due to the halogen being electropositive, leading to excitations that are thus mostly localized on the Cl (which is at

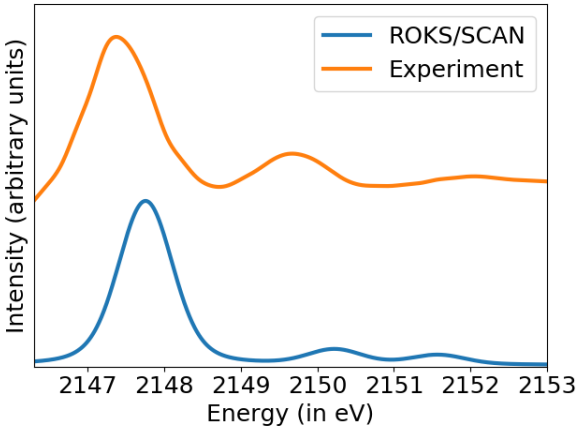
Species	Experiment	Vacuum			Solid		
		SCAN	SCANh	SCAN0	SCAN	SCANh	SCAN0
ClO_4^-	2835.1 [379]	2834.6	2834.9	2835.4	2834.7	2835.0	2835.4
ClO_3^-	2831.3 [379]	2830.8	2831.0	2831.4	2830.8	2831.1	2831.4
ClO_2^-	2826.8 [379]	2826.5	2826.7	2827.0	2826.6	2826.8	2827.1
$CuCl_4^{2-}$	2820.2 [380]	2820.1	2820.3	2820.6	2819.8	2819.9	2820.2
$NiCl_4^{2-}$	2821.5 [380]	2821.4	2821.6	2822.0	2821.1	2821.2	2821.6
$CoCl_4^{2-}$	2822.5 [380]	2822.3	2822.0	2822.5	2821.4	2822.4	2822.1
$FeCl_4^{2-}$	2822.8 [380]	2822.0	2822.3	2822.7	2821.7	2821.9	2822.3
$FeCl_4^-$	2820.5 [380]	2820.4	2820.4	2820.4	2820.2	2820.1	2820.1
RMSE		0.4	0.3	0.3	0.6	0.4	0.3
ME		-0.3	-0.2	0.2	-0.6	-0.3	-0.1
MAX		0.8	0.5	0.5	1.1	0.9	0.5

Table 4.4: Lowest dipole allowed Cl K-edge excitation energy for ionic species (in eV). All experimental data correspond to solid state measurements. ROKS was used for closed-shell species like ClO_4^-, and spin-unrestricted ΔSCF for open-shell systems like $CuCl_4^{2-}$. A mixed basis set (aug-pcX-2 on excitation site, aug-pcseg-1 on all other atoms) was utilized.

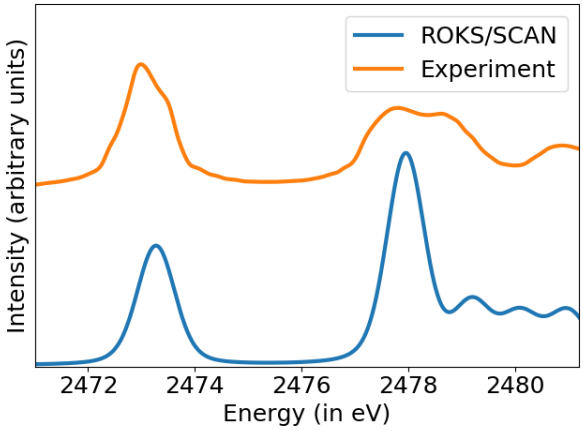
the center of the ion) and therefore reasonably isolated from the environment. Consequently, not much difference is found between predictions for vacuum, and an IEF-PCM model ionic solid. SCAN, SCANh and SCAN0 all fare reasonably at predicting excitation energies, with the former slightly underestimating and the latter slightly overestimating.

The second class of ions are high-spin tetrahedral transition metal chloride complexes where the lowest transition is charge-transfer (CT) from Cl to a singly occupied metal d level. In addition, the Cl site is on the periphery of the molecule, permitting greater influence from the environment. There is thus a perceptible red-shift in the IEF-PCM results relative to vacuum (due to greater stabilization of the CT like excited state). In addition, SCAN0 appears to be the best performer for this class of excitations. Nonetheless, SCANh appears to do a reasonable job at predicting excitation energies for all of the ionic systems, indicating that the OO-DFT/X2C approach remains capable of delivering semi-quantitative accuracy even outside of small molecules in the gas phase.

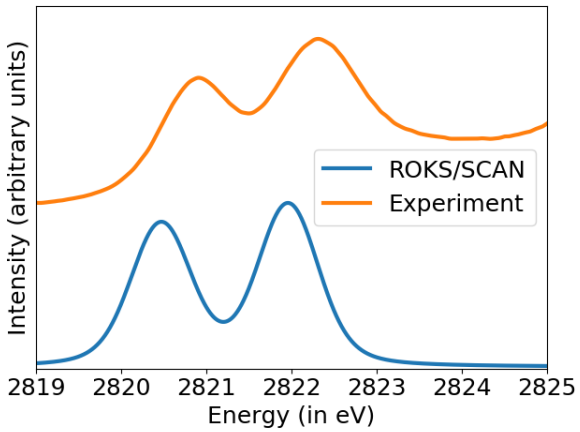
It is also instructive to look beyond computed excitation energies and consider the full spectrum. Fig 4.4 presents some examples where the experimental spectrum is compared to computed ones, without using any empirical translations. The peak energies align quite well, within the expected error range of ~ 0.5 eV. The peak heights agree less well, although in some cases this is clearly due to the experimental peaks having different widths (such as in Fig 4.4b for SO_2), vs the uniform broadening utilized for computed spectra. It would be interesting to compute spectral broadening directly from OO-DFT and determine if that leads



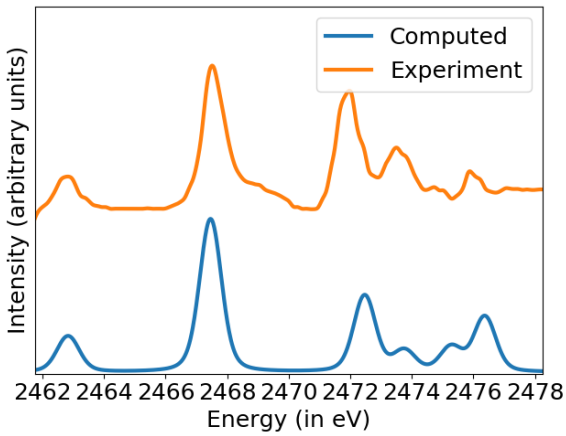
(a) P K-edge XAS of gaseous P_4O_6 [375], using a mixed basis (aug-pcX-2 on target P, aug-pcseg-1 on other atoms).



(b) S K-edge XAS of gaseous SO_2 [368], using a doubly augmented (d-aug-) pcX-2 basis on all atoms.



(c) Cl K-edge XAS of solid $[Ph_4P]_2TiCl_6$ [303]. The system was approximated with a $TiCl_6^{2-}$ ion, placed in a IEF-PCM dielectric utilizing NaCl parameters ($\epsilon_r = 6, n = 1.5$). A mixed basis (aug-pcX-2 on target Cl, aug-pcseg-1 on other atoms) was used.



(d) Combined S K-edge XAS (> 2470 eV) and S Kβ emission (< 2470 eV) of gaseous CSO [356]. ROKS was utilized for the XAS and unrestricted ΔSCF for the emission. A doubly augmented (d-aug-) pcX-2 basis on all atoms was used.

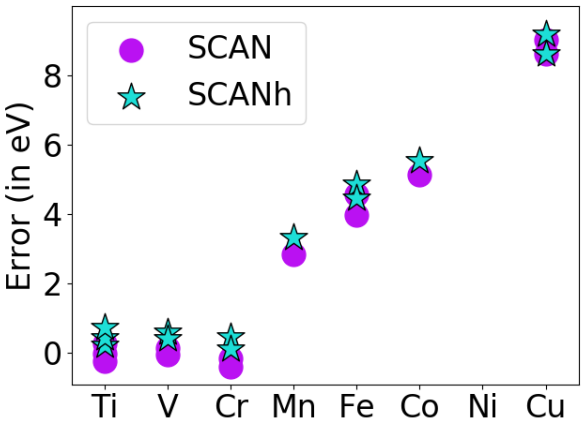
Figure 4.4: Comparison of experimental K-edge spectra for third period elements, and those computed with SCAN. Computed spectra were broadened with a Voigt profile with a Gaussian standard deviation of 0.3 eV and Lorentzian $\gamma = 0.121$ eV.

to better agreement between theory and experiment. We note that we computed intensities within the dipole-approximation (using transition dipole moments calculated in the manner described in the supporting information and utilizing non-orthogonal configuration interac-

tion techniques [381]). It is quite possible that higher order terms have a nonnegligible impact on the experimental X-ray spectrum [382]. Nonetheless, it appears that OO-DFT/X2C is quite effective at reproducing experimental spectra for third period elements.

	Experiment		SCAN	SCANh
$TiCl_4$	4969.2	[378]	4968.9	4969.4
$TiCpCl_3$	4968.1	[378]	4968.1	4968.5
$TiCp_2Cl_2$	4967.3	[378]	4967.5	4968.0
$VO(acac)_2$	5468.4	[383]	5468.6	5469.0
VCp_2Cl_2	5468.4	[383]	5468.3	5468.8
CrO_4^{2-}	5996.5	[384]	5996.3	5996.9
$Cr_2O_7^{2-}$	5996.6	[384]	5996.2	5996.7
MnO_4^-	6543.3	[385]	6546.1	6546.6
$FeCp_2$	7111.9	[386]	7116.5	7116.8
$FeCl_4^-$	7113.2	[285]	7117.1	7117.6
$CoCl_4^{2-}$	7709.2	[387]	7714.3	7714.7
$CuCl_4^{2-}$	8977.6	[388]	8986.6	8986.8
$Cu(CF_3)_4^-$	8981.8	[388]	8990.4	8990.4

(a) Individual values (in eV)



(b) Excitation energy error vs nuclear charge.

Table 4.5: Lowest symmetry allowed transition metal K-edge transitions. ΔSCF was used for 1s→SOMO transitions of open-shell systems like $VO(acac)_2$, while ROKS was used for closed-shell species like $TiCl_4$. A mixed basis (decontracted aug-cc-pωCVTZ [389] on excitation site, aug-cc-pVDZ on all other atoms) was utilized. Experimental data corresponds to the solid state for all species other than $TiCl_4$, whose spectrum was collected in toluene solution.

We next shift our attention to the 3d transition metals. Transition metal complexes are often open-shell, and have several low-lying orbitals that core electrons can be excited to. K-edge spectra of these species are therefore widely studied, despite the 1s→3d transition in bare atoms being dipole forbidden. Indeed, the transition in centrosymmetric entities (like octahedral complexes) can only be driven by electric quadrupole terms or vibrational symmetry breaking [285, 390]. Tetrahedral complexes however can have some p-d mixing, leading to some dipole allowed intensity for K-edge transitions [390]. Furthermore, X-ray emission spectra (XES) can be collected for $2p/3p \rightarrow 1s$ dexcitations, yielding further useful information [284].

Table 4.5 shows the performance of OO-DFT in predicting transition metal K-edge energies, with the SCAN and SCANh functionals. The results are quite adequate for Ti, V, and Cr but performance is significantly degraded for heavier transition metals like Mn, Fe, Co and Cu due to significant overestimation (as made evident by Fig 4.5b). This is quite interesting, as SCAN had a slight penchant for underestimation when it came to lighter elements, unlike the significant overbinding observed here for Cu, Co and Fe (beyond the scale

of typical OO-DFT errors). It should be noted that many of the species in Table 4.5 are high-spin tetrahedral complexes like $FeCl_4^-$,$CoCl_4^{2-}$ and $CuCl_4^{2-}$ or d^0 species like $TiCl_4$, CrO_4^{2-} and MnO_4^-. These species are therefore not particularly multireference, removing one possible source of the error. While it is possible that this error arises from inadequacies of the density functional approximation, the sudden and rapid increase in error with increasing Z seems to suggest that relativistic effects are playing a significant role.

Indeed, the spin-free one-electron X2C model employed in our work is hardly complete, as it does not include vector (spin-orbit) terms, as well as any effects of relativity on the two-electron terms. The spin-orbit terms are unlikely to be relevant for K-edges, but the missing two-electron terms are of a comparable magnitude to the error. Specifically, X2C does not transform the electron-electron repulsion terms (the so-called "picture-change" [128]), incorporate additional (Breit) [120] terms in the electron-electron interaction terms, or account for quantum electrodynamic effects. These additional terms sum to ~ 20 eV of underbinding for the Kr K-edge [341, 391, 392], indicating that overestimation by a few eV is quite possible for heavier fourth period elements from lack of such effects alone. In fact, the good performance of SCAN up to Cr is possibly fortuitous to some degree. However, it is worth stressing that the X2C model nonetheless manages to account for the vast majority of relativistic effects for even species like Cu/Kr. It is also worth noting that translating OO-DFT/X2C metal K-edge spectra for alignment with experiment would involve much smaller shifts than TDDFT, reducing the magnitude of potential translation driven error. Nonetheless, we wish to avoid any need for empirical translation of spectra, and instead intend to pursue more accurate relativistic models to quantitatively model the K-edge spectra of heavier elements.

	Experiment			Theoretical L_3 edge		
	L_3	L_2	3J=L_2-L_3	SCAN	SCANh	SCAN0
$TiCl_4$	456.9 [393]	462.5	5.6	455.9	456.4	457.1
$Mn(OH_2)_6^{2+}$	639.7 [394]	649.1	9.4	638.7	638.9	639.3
$Fe(CN)_6^{3-}$	705.8 [395]	718.4	12.6	706.1	706.5	707.1
$FeCp_2$	708.9 [396]	721.2	12.3	708.2	708.5	709.0
$CuCl_4^{2-}$ (D_{2h})	930.1 [388]	950.1	20	929.4	929.5	929.8
$Cu(CF_3)_4^-$	934.7 [388]	954.7	20	933.5	933.5	933.7

Table 4.6: Lowest symmetry allowed L-edge (2p) transitions for transition metal containing species (in eV). ΔSCF was used for 2p$\rightarrow$SOMO transitions of open-shell systems like $CuCl_4^{2-}$, while ROKS was used for closed-shell species like ferrocene ($FeCp_2$). Computed multiplet averaged energies (found from averaging over all the 2p orbitals) were red-shifted by J (where $3J$ is the experimental energy gap between the L_2 and L_3 peaks) to better approximate the experimental L_3 peaks. A mixed basis (decontracted aug-cc-pωCVTZ on excitation site, aug-cc-pVDZ on all other atoms) was utilized for the calculations.

Metal L-edge (2p) spectra are also extensively studied via experiment. Accurate com-

putation of spectra necessitates going beyond the scalar relativistic paradigm as the degeneracy of the $2p$ levels is broken by spin-orbit coupling. Scalar relativistic models like X2C can potentially yield a reasonable estimate for the multiplet averaged peak, which can be subsequently shifted by atomic spin-orbit values to obtain the experimentally observed $L_{3/2}$ (L_3-edge) and $L_{1/2}$ (L_2-edge) multiplet peaks. This strategy has been found to be effective for L-edges of Si,P,S and Cl [321]. However, it is less appealing for heavier elements due to the larger magnitude of the multiplet splitting (20 eV for Cu vs 1.6 eV for Cl), which can potentially exert a direct influence in the OO process. We have nonetheless applied this measure to compute L-edges for a few 3d metal containing species (shown in Table 4.6), to assess this approach. It is quite apparent that the errors are much larger here (relative to Fig 4.3), with significant underestimation being the norm for all species aside for $[Fe(CN)_6]^{3+}$. The results are also often quite sensitive to % HF exchange (more so than most of the species considered till this point). Nonetheless, the worst case errors are just slightly above 1 eV, indicating that OO-DFT with proper inclusion of spin-orbit effects and careful functional choice has the potential to be quite accurate in predicting metal L-edges. We are presently investigating this aspect further.

4.5 Conclusions

Overall, it is clear that the OO-DFT/X2C combination is capable of consistently delivering $<$ 1 eV error for core-level excitation/ionization energies for third period elements and reproduce experimental spectra fairly well. Indeed, the typical error is $\sim$ 0.5 eV for SCAN/SCANh, as shown by Fig 4.3 and the RMSEs reported in Tables 4.1 and 4.2. X2C therefore extends the applicability of OO-DFT methods to elements heavier than F, which was previously the limit for computational core-level spectroscopy with such methods (without post-facto application of ad-hoc relativistic corrections). OO-DFT X2C also appears to be adequate for the K-edges of the transition metals Ti, V and Cr. However, performance degrades starting with Mn, possibly due to lack of relativistic effects in the two-electron interaction terms. The spin-free X2C model is also incapable of accounting for the spin-orbit splitting observed in experimental L-edge spectra. Accounting for these effects would be critical for extension of OO-DFT beyond the cases explored in this work. We are also attempting to apply OO-DFT extensively to K-edge spectra of third period elements, in order to uncover any additional limitations of the approach. Any resulting insight could also prove valuable in training density functionals that are accurate in modeling both ground states and OO-DFT excited states, as it is quite possible that not much further improvement in prediction quality can be obtained from functionals trained solely for the ground state.

Supplementary Material

Supplementary information for this work can be accessed via Ref. 397. They include additional X2C derivations, a short note on S K-edge binding energies of some species, raw data and geometries of all species investigated in this work.

Chapter 5

Electron-Affinity Time-Dependent Density Functional Theory

5.1 Introduction

Recent advancements in synchrotron and ultrafast tabletop X-ray light-sources mark the dawn of an X-ray technology renaissance. With exceptional element-specificity, X-ray spectroscopy has found use in probing liquid-to-metal phase transitions of ammonia [398], tracking charge-separation dynamics in dye-sensitized solar cells and organic light-harvesting systems [399, 400], and has revealed quantum nuclear dynamics near conical intersections [401]. Modern X-ray absorption spectroscopy (XAS) is capable of energy resolution on the order of 0.2–0.9 eV [402], which is well below the error statistics of most modern theoretical methods that are routinely used to model XAS.

Linear-response time-dependent density functional theory (TDDFT) is by far the most commonly used method for computing excitation energies due its accuracy and efficiency [247, 248, 403–406]. While formally exact for excitation energies, TDDFT in practice is approximate due to inexact ground state functionals and the ubiquitous adiabatic approximation (henceforth assumed) [92, 248, 407]. Although TDDFT achieves statistical accuracy of $\sim$0.2–0.3 eV for valence excitations [408], errors increase dramatically for core excitations, often requiring empirical shifts on the order of 10–100 eV to realign the calculated spectra with experiment [329, 409–411]. While range-separated hybrid functionals perform better in this regard [412, 413], specialized short-range corrected (SRC) functionals that feature a large amount of short-range Hartree-Fock (HF) exchange to correct for differential self-interaction error in the core have also been used instead of empirical shifting, albeit to the disregard of broader thermochemical properties [296, 298–300, 414–417]. Apart from pure TDDFT, semi-empirical extensions of configuration interaction that employ Kohn-Sham orbitals have been applied with some success to core-excitations [418–420]. In some cases, particularly in periodic systems, TDDFT and configuration-interaction methods are sidestepped in favor of cruder approaches like the Slater transition or transition potential methods [421–429].

Despite the myriad ways in which XAS can be calculated, in this work our focal point is linear-response TDDFT.

One source of error in predicting core-excitations using linear-response theory is the large orbital relaxation effect that follows from the displacement of charge out of a core orbital [430, 431]. This can be addressed on a state-by-state basis using orbital-optimized density functional theory (OO-DFT), which explicitly relaxes the orbitals of excited-state configurations [238]. While OO-DFT routinely achieves a statistical accuracy of $\sim$0.3 eV for core-excitations [321, 397], state-by-state optimization is far less efficient than full-spectrum methods like TDDFT. OO-DFT also requires some *a priori* knowledge of the system, complicating the selection of the "correct" set of bespoke determinants in systems with a high density of states.

Orbital relaxation error is related to the fundamentally incorrect particle-hole interaction in TDDFT descriptions of core-excited states [91, 282, 432]. This is the major source of error in TDDFT; emerging from the fact that the virtual orbitals in DFT are optimized in the n-electron potential, causing incomplete cancellation of the interaction of the excited electron with itself in the (previously occupied) core orbital. For example, consider the pure particle-hole interaction that results from exciting an electron between two molecular orbitals (MOs) that have zero overlap. For global hybrid functionals, the only nonzero elements of the orbital Hessians belong to the **A** matrix,

$$A_{ia,jb} = (\varepsilon_a - \varepsilon_i)\delta_{ij}\delta_{ab} - C_{\mathrm{HF}}(ij|ab) \tag{5.1}$$

where C_{HF} is the coefficient of HF exchange. The Coulomb interaction $(aa|ii)$ is included in the orbital energy difference and only in the case of exact exchange ($C_{\mathrm{HF}} = 1$) is this interaction properly cancelled by the third term in Eq. 5.1, leading to particle-hole attraction. Therefore with approximate density functionals, the excited electron "feels" a residual Coulomb potential from its unexcited image rather than a proper particle-hole attraction, causing core-excitation energies to be dramatically underestimated in approximate TDDFT.

In this work, we introduce a linear-response TDDFT formalism that effectively models particle creation in the virtual space from an n–1-electron reference density. This way, orbital relaxation and information about the core hole are built directly into the reference density, completely eliminating electron-hole self-interaction error (eh-SIE) by construction. Our method generalizes the static-exchange approximation (STEX) into a density functional theory (DFT) framework.

Such generalizations have recently been proposed based purely on error cancellation between restricted open-shell Kohn-Sham (ROKS) theory and STEX [433], but this work aims at a fully derivable formalism. Herein we demonstrate multifaceted benefits of an exact approach, including better overall performance, and a recovery of the Jacob's Ladder concept in DFT. While TDDFT is the workhorse of excited-state calculations in quantum chemistry, we further note that the concept of adding electrons to an n–1-electron reference determinant has been employed within Green's function based GW methods [434, 435] and within algebraic diagrammatic construction approaches [436, 437] to account for orbital relaxation

and (in the case of GW theory) for a more appropriate description of particle-hole interactions. In principle, our proposed approach recovers the same poles as the single-particle Green's function in the electron addition domain. However, unlike GW approaches that scale roughly as $\mathcal{O}(N^4)$ with a nontrivial prefactor (where N is the number of basis functions) [438], our proposed approach has the same scaling as STEX ($\mathcal{O}[V^3]$, where V is the number of unoccupied MOs), making this a far more appealing method for large systems.

5.2 Theoretical Background

5.2.1 Static-Exchange Approach for Core-Excited States

The STEX formalism has been used to improve upon core-excitation energies offered by configuration interaction with single excitations (CIS) [91, 439, 440] for a number of years [441, 442]. In brief, STEX involves optimizing the MOs of the n–1-electron (core-ionized) system, followed by an electron-affinity CIS (EA-CIS) calculation to reattach the missing electron to the virtual orbitals, thereby yielding a partially orbital-optimized core-excitation spectrum that accounts for the strong polarization effect from creating a core hole. For a closed-shell reference, the EA-CIS equations for singlet states take the form,

$$A_{ia,ib} = E_{\mathrm{HF}}^{(n-1)}\delta_{ab} + F_{ab}^{(n-1)} + (ia|ib) \;, \tag{5.2}$$

where i is the core hole MO, $E_{\mathrm{HF}}^{(n-1)}$ is the core-ionized reference energy, $F_{ab}^{(n-1)}$ are elements of the virtual-virtual block of the core-ionized Fock matrix, and $(ia|ib)$ is an exchange integral in the standard Mulliken notation. Diagonalizing $\mathbf{A}$ results in states that are orthogonal to the core-ionized reference determinant, but are not orthogonal to the original n-electron ground state. The final step of the STEX procedure involves constructing nonorthogonal configuration interaction (NOCI) elements to project the n-electron ground state out of the Hamiltonian prior to diagonalizing, ensuring that all excited states are strictly orthogonal to the initial n-electron ground state determinant [443–445]. Herein, we will show that the nonorthogonality of the excited states to the ground state can be safely ignored when calculating K-edge XAS with almost no impact on the predicted excitation energies or transition properties, thus paving the way for a TDDFT formalism where the ambiguity of DFT-based NOCI elements once hindered such developments.

5.2.2 EA-TDDFT: generalizing STEX to a DFT formalism

In order to generalize STEX to a TDDFT framework, we will use continuum MOs as a derivation tool. For our purposes, continuum MOs are fictitious, ultra-diffuse orbitals that do not interact with other MOs in the system and have zero energy. They offer utility in derivations of particle-nonconserving processes by recasting particle creation/annihilation into the language of particle-conserving excitations [446].Throughout this work, we reserve the labels $\{j, k, l, \dots\}$ to denote occupied MOs, $\{a, b, c, \dots\}$ for the virtual MOs, and $\{p, q, r, \dots\}$ refer

to general orbitals. Specific notation is reserved for the continuum MO, designated as x, and the core-hole MO, i.

To ameliorate orbital relaxation error, we begin with the self-consistently optimized MOs for the core-ionized system. We are interested in a protocol that uses particle-conserving excitations that emulate the action of the particle creation operator on our core-ionized reference,

$$|\Psi_i^a\rangle = \hat{a}_a^\dagger |\Psi_0\rangle \ , \tag{5.3}$$

where $|\Psi_0\rangle$ is the core-ionized reference determinant, and $\hat{a}_a^\dagger$ is the creation operator. One possibility that retains correlations between single excitations in the response theory that follows is to consider two successive excitations $x \to i$ and $i \to a$ out of a modified core-ionized reference determinant that includes a single continuum MO. Conceptually, this can be likened to excited-state absorption where the n-electron state with the core-hole MO reoccupied acts as the intermediate state. In operator form it can be readily shown that,

$$|\Psi_i^a\rangle = \hat{a}_a^\dagger \hat{a}_i \hat{a}_i^\dagger \hat{a}_x |\Psi_0 \chi_x\rangle = \hat{a}_a^\dagger \hat{a}_i \hat{a}_i^\dagger |\Psi_0\rangle = \hat{a}_a^\dagger |\Psi_0\rangle \ , \tag{5.4}$$

where $|\Psi_0 \chi_x\rangle$ is the modified core-ionized reference, containing the noninteracting spin-orbital χ_x. This exercise reveals that the successive particle-conserving excitations $x \to i$ and $i \to a$ indeed reduce to particle creation in orbital a of the *unmodified* core-ionized reference determinant, which itself can be viewed as the tensor product of the core-ionized reference with the vacuum level in the space of continuum orbitals.

In order to capture this process in the language of density matrices, such that our protocol is amenable to DFT, we consider two successive linear responses. The first response generates the n-electron density from the n–1-electron reference by exciting an electron from a continuum MO into the core hole, and the second response yields eh-SIE-corrected excitations of this (newly added) core electron into the virtual space. Throughout this derivation, we follow the density matrix formalism starting from the Liouville-von Neumann equation [91, 404],

$$i\frac{\partial \mathbf{P}(t)}{\partial t} = [\mathbf{F}(t), \mathbf{P}(t)] \ . \tag{5.5}$$

The first response is obtained by restricting the excitation space to the (occupied) continuum MO and the (unoccupied) core-hole to yield,

$$\begin{aligned} A_{xi,xi} &= F_{ii}^{(n-1)} \\ B_{xi,xi} &= 0 \end{aligned} \tag{5.6}$$

Because the continuum MO does not interact with the rest of the system, all two-electron integrals involving the continuum MO vanish to give an expression that corresponds to the negative electron affinity in the limit of the exact functional [447]. Importantly, no orbital rotations are encoded in this response, meaning that the (idempotent) n-electron density can be exactly constructed with the n–1-electron MOs to linear order.

At some time $t' > t$, we apply a second time-varying electric field to the perturbed n-electron system. By regenerating the n-electron system, we have reintroduced eh-SIE, so we now seek to separate the response due to the presence of the core electron from the remainder of the response to the applied field. Assuming that the n-electron density is not too far from a stationary point, we may write the perturbed density and corresponding Fock matrices as,

$$\begin{aligned}\mathbf{P}(t') &= \mathbf{P}_0^{(n)} + \delta\mathbf{P}_{\mathrm{CO}}(t') + \delta\mathbf{P}_{\mathrm{EF}}(t') \\ \mathbf{F}(t') &= \mathbf{F}_0^{(n)} + \delta\mathbf{F}_{\mathrm{CO}}(t') + \delta\mathbf{F}_{\mathrm{EF}}(t')\end{aligned} \tag{5.7}$$

where $\mathbf{P}_0^{(n)}$ is the static part of the n-electron density, $\delta\mathbf{P}_{\mathrm{CO}}(t')$ represents the component of the response due to the (now occupied) core MO and $\delta\mathbf{P}_{\mathrm{EF}}(t')$ indicates the response of the n-electron system to the second electric field. Substituting Eq. 5.7 into Eq. 5.5 and keeping the terms that are linear with respect to the perturbing field leads to,

$$\begin{aligned}i\frac{\partial\delta\mathbf{P}_{\mathrm{CO}}(t')}{\partial t'} + i\frac{\partial\delta\mathbf{P}_{\mathrm{E}F}(t')}{\partial t'} &= [\mathbf{F}_0^{(n)}, \delta\mathbf{P}_{\mathrm{CO}}(t')] \\ + [\delta\mathbf{F}_{\mathrm{CO}}(t'), \mathbf{P}_0^{(n)}] + [\mathbf{F}_0^{(n)}, \delta\mathbf{P}_{\mathrm{E}F}(t')] &+ [\delta\mathbf{F}_{\mathrm{E}F}(t'), \mathbf{P}_0^{(n)}]\end{aligned}, \tag{5.8}$$

which is simply the sum of two linear responses. Notably, this formalism has been used to subtract nonstationary oscillations out of real-time TDDFT simulations of excited-state absorption (including application to transient XAS), and we have adopted similar notation throughout [448–452].

We use the fact that the above responses are uncoupled to correct the n-electron response by subtracting the components that emerge due to the occupied core orbital *via* the difference Fock matrix,

$$\begin{aligned}F_{pq}^{\mathrm{CO}} &= F_{pq}^{(n)} - F_{pq}^{(n-1)} \\ &= (ii|pq) - C_{\mathrm{HF}}(ip|iq) \\ &\quad + (1 - C_{\mathrm{HF}})(p|V_{\mathrm{xc}}^{(n)} - V_{\mathrm{xc}}^{(n-1)}|q)\end{aligned} \tag{5.9}$$

and its corresponding density (all n-electron quantities are constructed using the n–1-electron MOs). This form of the Fock matrix incorporates all zeroth-order couplings between n- and n–1-electron potentials without approximation, and the corresponding difference density is idempotent with one electron in core MO i, permitting excitations of the form $i \to a$.

Subtracting the response of the core orbital density from that of the n-electron density (see Sec. 5.2.5 for details), leads to a eh-SIE-corrected n-electron response in terms of n–1-electron quantities

$$\begin{aligned}A_{ia,ib} &= F_{ab}^{(n-1)} - F_{ii}^{(n-1)}\delta_{ab} \\ &\quad + (ia|ib) + (1 - C_{\mathrm{HF}})(ia|f_{\mathrm{xc}}^{(n-1)}|ib) \\ B_{ia,ib} &= (ia|ib) + (1 - C_{\mathrm{HF}})(ia|f_{\mathrm{xc}}^{(n-1)}|ib)\end{aligned} \tag{5.10}$$

where,

$$f_{\mathrm{xc}}^{(n-1)} = \frac{\partial V_{\mathrm{xc}}[\rho^{(n-1)}]}{\partial \rho^{(n-1)}} \,. \tag{5.11}$$

Finally, we add this to the result of the initial response in Eq. 5.6 to obtain the working equations,

$$\begin{aligned} A_{ia,ib} &= F_{ab}^{(n-1)} + (ia|ib) + (1 - C_{\mathrm{HF}})(ia|f_{\mathrm{xc}}^{(n-1)}|ib) \\ B_{ia,ib} &= (ia|ib) + (1 - C_{\mathrm{HF}})(ia|f_{\mathrm{xc}}^{(n-1)}|ib) \end{aligned} \tag{5.12}$$

Each of the above responses comes from exact TDDFT, suggesting that with time-dependent exchange-correlation kernels this approach could be made exact (to first order). Of course, knowledge of the exact functional would render this formalism obsolete because the exact functional is asymptotically correct (eh-SIE-free) and has frequency dependence (accounting for orbital relaxation) [453]. From a utilitarian perspective, the exact functional is not available and all practical TDDFT implementations employ the adiabatic local density approximation (ALDA). Therefore, Eq. 5.12 can be viewed as a pragmatic correction to errors associated with the ALDA in TDDFT for XAS.

By nature of the core-ionized reference determinant and because the MOs do not relax on addition of the electron, the orbital relaxation codified into the n–1-electron density is retained. The second response is also eh-SIE-corrected, ensuring that there is no residual Coulomb-like interaction between excited electron and core hole. In fact, our proposed correction (Eq. 5.19) bears a delightful resemblance to the virtual-orbital self-interaction definition proposed by Imamura and Nakai [454], with the added benefit that our equations capture orbital relaxation. This immediately suggests a metric for quantifying the extent of eh-SIE *via* the eigenvalues of the core-orbital response matrix (Eq. 5.21).

In the limit of the Hartree-Fock functional ($C_{\mathrm{HF}} = 1$) this metric is exactly zero, and under the Tamm-Dancoff approximation (TDA) [258] Eq. 5.12 becomes precisely equivalent to the EA-CIS equation (Eq. 5.2). This implies that Eq. 5.12 is a generalization of EA-CIS to a DFT framework, so we call our approach electron-affinity TDDFT (EA-TDDFT).

In true analogy to EA-CIS, we will employ the TDA (EA-TDA) throughout this work, setting the $\mathbf{B}$ matrix to zero in Eq. 5.12. While in some applications the modified reference state (*e.g.* a core-ionized determinant) can lead to difficulties in solving the full non-Hermitian eigenvalue problem due to orbital rotations that drive the solution towards the ground state [455–458], we have found that such problems are not encountered in EA-TDDFT[459]. Despite the fact that the full EA-TDDFT equations can be readily solved, the TDA is likely an excellent approximation within the confines of core excitations associated with K-edge XAS, as the elements of $\mathbf{B}$ are quite small. For the sake of comparison, EA-TDA is also a more direct analogue to STEX (a CIS theory), which is what we are attempting to generalize to DFT.

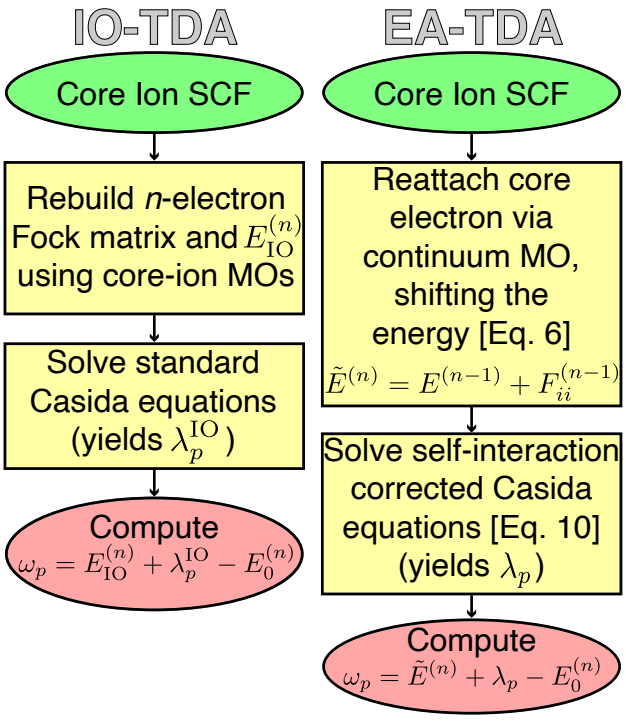


Figure 5.1: Flowchart describing the conceptual protocol for (left) an IO-TDA calculation and (right) an EA-TDA calculation, where $E_{\mathrm{IO}}^{(n)}$ is the energy of the n-electron Fock matrix constructed *via* core-ion orbitals, $E_0^{(n)}$ is the SCF ground-state energy, λ_p^{IO} are the eigenvalues of the Casida equations using the IO-TDA reference, λ_p are the eh-SIE corrected eigenvalues of EA-TDA, and ω_p are the core-excitation energies.

5.2.3 Shortcomings of the Ion-Orbital TDDFT Approach

Apart from EA-TDA, we also consider the more naive approach of optimizing the orbitals of the n–1-electron system and using them directly to reconstruct the n-electron density. From this nonstationary initial state, we perform TDDFT under the TDA using the usual Casida formulation. In principle, this ion-orbital TDA (IO-TDA) approach incorporates orbital relaxation into the reference but lacks the ingredients that account for eh-SIE (details in Section 5.2.4).

It is important to note that EA-TDA differs strongly from IO-TDA in two respects. First, the energy of the intermediate n-electron state is constructed differently between the two methods. Whereas IO-TDA forms $E_{\mathrm{IO}}^{(n)}$, the energy of the n-electron Fock matrix using the orbitals of the core-ion with the core electron reattached, EA-TDA constructs $\tilde{E}^{(n)} = E^{(n-1)} + F_{ii}^{(n-1)}$. The two energies are only equivalent in the case of HF where the Fock matrix is strictly linear in the density. Second, we emphasize that Eq. 5.10 is not just a standard linear-response TDDFT expression, but one that encodes an exact first-

order eh-SIE correction to the TDDFT equations. Without this correction, EA-TDA would correct for orbital relaxation but would not correct for the (even larger) self-interaction error in the excited state. The flowchart in Fig. 5.1 emphasizes the difference between the eigenvalues λ_p^{IO} of IO-TDA, which corrects for orbital relaxation but not eh-SIE, and the self-interaction corrected eigenvalues, λ_p, of EA-TDA. Later in this work, we investigate the impact of orbital relaxation and eh-SIE, showing that both must be adequately compensated for accurate results. Finally, we note that explicitly performing each step in the conceptual protocol in Fig. 5.1 is not necessary in practice, where Eq. 5.12 can be directly constructed and diagonalized to yield electron affinities γ_p such that $\omega_p = E^{(n-1)} + \gamma_p - E_0^{(n)}$.

This definition of the excitation energies is common between EA-TDA and STEX, but unlike STEX, EA-TDA does not use a projection operator to ensure orthogonality between the singly-excited configurations and the n-electron ground state. Nonorthogonality can have detrimental effects on transition dipole moments (TDMs) even when energies are largely unaffected [460], so we take it into account by using a pseudo-wavefunction *ansatz* (as done in TDDFT) to compute an overlap-free TDM (see Sec. 5.2.7 for details). We justify this approach by comparison of EA-TDA with the HF functional against STEX on a data set of 132 experimental K-edge transitions of small molecules. Our results (see Tables S1–S3 in the SI) reveal strong agreement between EA-TDA(HF) and STEX, with a maximum difference in transition energies of just 0.1 eV and a mean difference of 0.01 eV. The average difference between EA-TDA(HF) and STEX oscillator strengths is $\sim 10^{-5}$, with a maximum difference of $\sim 10^{-4}$ while the average STEX oscillator strength is 10^{-2}. We therefore conclude that the nonorthogonality of the final states in EA-TDA exerts a minimal influence on the details of the spectrum.

5.2.4 Linear-Response Time-Dependent Density Functional Theory and its Ion-Orbital Variant

The standard time-dependent density functional theory (TDDFT) orbital Hessians are used for the "ion-orbital" TDDFT approach, albeit from a nonstationary n-electron reference state that is constructed from the $n-1$-electron molecular orbitals (MOs) of the core-ionized system. The usual TDDFT **A** and **B** matrices take the form,

$$\begin{aligned} A_{ia,ib} &= E^{(n)}\delta_{ab} + F_{ab}^{(n)} - \varepsilon_i^{(n)}\delta_{ab} + (ia|ib) - C_{\mathrm{HF}}(ii|ab) + (1 - C_{\mathrm{HF}})(ia|f_{\mathrm{xc}}^{(n)}|ib) \\ B_{ia,ib} &= (ia|ib) - C_{\mathrm{HF}}(ib|ai) + (1 - C_{\mathrm{HF}})(ib|f_{\mathrm{xc}}^{(n)}|ai) \end{aligned} \tag{5.13}$$

where $f_{\mathrm{xc}}^{(n)}$ is the exchange-correlation kernel, defined as,

$$f_{\mathrm{xc}}^{(n)} = \frac{\partial V_{\mathrm{xc}}[\rho^{(n)}]}{\partial \rho^{(n)}} \,, \tag{5.14}$$

and where all quantities denoted with superscript (n) are computed using the n-electron density. In the case of IO-TDDFT, these n-electron quantities are constructed from the

n-electron density built from the unrelaxed $n-1$-electron MOs of the core-ionized system:

$$P_{\mu\nu}^{(n)} = \sum_{i}^{N} C_{\mu i}^{(n-1)} (C_{\nu i}^{(n-1)})^{*} \tag{5.15}$$

5.2.5 Derivation of the $n-1$-electron Response Kernel

In order to correct for the particle-hole interaction error encountered in the intermediate n-electron state obtained after electron addition from the continuum MO, we take the response of the applied field on the n-electron state, which yields the Casida equations for the restricted case,

$$\begin{aligned} A_{ia,ib}^{(n)} &= E^{(n)}\delta_{ab} + F_{ab}^{(n)} - F_{ii}^{(n)}\delta_{ab} + 2(ia|ib) - C_{\mathrm{HF}}(ii|ab) + (1-C_{\mathrm{HF}})(ia|f_{\mathrm{xc}}^{(n)}|ib) \\ B_{ia,ib}^{(n)} &= 2(ia|ib) - C_{\mathrm{HF}}(ib|ai) + (1-C_{\mathrm{HF}})(ib|f_{\mathrm{xc}}^{(n)}|ai) \end{aligned} \tag{5.16}$$

and subtract the response of the core orbital with associated the Fock matrix elements,

$$F_{pq}^{\mathrm{CO}} = F_{pq}^{(n)} - F_{pq}^{(n-1)} = (ii|pq) - C_{\mathrm{HF}}(ip|iq) + (1-C_{\mathrm{HF}})(p|V_{\mathrm{xc}}^{(n)} - V_{\mathrm{xc}}^{(n-1)}|q) \ , \tag{5.17}$$

where $F_{pq}^{(i)}$ is the core electron's contribution to the Fock matrix of the n-electron system. This form of the Fock matrix accounts for all couplings between the core-electron components and the remainder of the n-electron density. The associated density matrix is idempotent and contains one electron in the core orbital, naturally constraining the excitations *via* the idemptency condition such that they can only emerge from core MO i. The response for the corresponding density matrix takes the form,

$$\begin{aligned} A_{ia,ib}^{\mathrm{CO}} &= E^{\mathrm{CO}}\delta_{ab} + F_{ab}^{\mathrm{CO}} - F_{ii}^{\mathrm{CO}}\delta_{ab} + \frac{\partial \mathbf{F}_{ia}^{\mathrm{CO}}}{\partial \mathbf{P}_{ib}} \\ B_{ia,ib}^{\mathrm{CO}} &= \frac{\partial \mathbf{F}_{ai}^{\mathrm{CO}}}{\partial \mathbf{P}_{ib}} \end{aligned} \ , \tag{5.18}$$

where $E^{\mathrm{CO}} = \tilde{E}(n) - E_0(n-1)$ (the nonstationary n-electron energy minus the stationary $n-1$-electron energy of the core ion) and the partial derivatives yield the final expression for the core-orbital response,

$$\begin{aligned} A_{ia,ib}^{\mathrm{CO}} &= E^{\mathrm{CO}}\delta_{ab} + F_{ab}^{\mathrm{CO}} - F_{ii}^{\mathrm{CO}}\delta_{ab} + (ia|ib) - C_{\mathrm{HF}}(ii|ab) + (1-C_{\mathrm{HF}})(ia|f_{\mathrm{xc}}^{(n)} - f_{\mathrm{xc}}^{(n-1)}|ib) \\ B_{ia,ib}^{\mathrm{CO}} &= (ia|ib) + (1-C_{\mathrm{HF}})(ia|f_{\mathrm{xc}}^{(n)} - f_{\mathrm{xc}}^{(n-1)}|ib) \end{aligned} \ . \tag{5.19}$$

Finally, subtracting the core-orbital part of the response from the full n-electron response leads to,

$$\begin{aligned} A_{ia,ib}^{(n)} - A_{ia,ib}^{\mathrm{CO}} &= E_0(n-1)\delta_{ab} + F_{ab}^{(n-1)} - F_{ii}^{(n-1)}\delta_{ab} + (ia|ib) + (1-C_{\mathrm{HF}})(ia|f_{\mathrm{xc}}^{(n-1)}|ib) \\ B_{ia,ib}^{(n)} - B_{ia,ib}^{\mathrm{CO}} &= (ia|ib) + (1-C_{\mathrm{HF}})(ia|f_{\mathrm{xc}}^{(n-1)}|ib) \end{aligned} \ . \tag{5.20}$$

We note here that the energy E^{CO} is equal to the energy of orbital i only for the exact functional or Hartree-Fock theory, so the explicit form of this energy is never assumed.

5.2.6 Long-Range Self-Interaction Metric

Within an ion-orbital *ansatz* such as IO-TDA, Eq. 5.19 is suggestive of a metric that can be used to quantify the degree of long-range self-interaction error (*i.e.* the degree of inexact particle-hole interaction) in approximate density functionals. Considering only the change in the excitation energy offered by the core-orbital correction, the total particle-hole interaction error for TDA approximations is,

$$A^{\mathrm{CO}}_{ia,ib} = F^{\mathrm{CO}}_{ab} - F^{\mathrm{CO}}_{ii}\delta_{ab} + (ia|ib) - C_{\mathrm{HF}}(ii|ab) + (1 - C_{\mathrm{HF}})(ia|f^{(n)}_{\mathrm{xc}} - f^{(n-1)}_{\mathrm{xc}}|ib) \tag{5.21}$$

In Hartree-Fock theory Eq. 5.17 implies that $F^{\mathrm{CO}}_{ii} = 0$ and that $F^{\mathrm{CO}}_{ab} + (ia|ib) - (ii|ab) = 0$, resulting in a long-range self-interaction error of exactly zero. It also implies that IO-TDA with the HF functional should give equivalent results to STEX if the non-orthogonality with the n-electron ground state is not projected out of the STEX Hamiltonian. This is indeed the case, as IO-TDA and EA-TDA produce exactly the same results if the HF functional is used. If this metric produces a nonzero value, then the density functional approximation being used incurs some degree of inexact particle-hole interaction and the larger the value of the metric, the larger the long-range self-interaction error of the functional.

5.2.7 Overlap-Free Transition Dipole Moments

The EA-TDA spectrum is comprised of states, $\{\Psi^a_i\}$, that are not orthogonal to the ground state reference, Φ_0, which must be considered when computing transition properties. Despite our double-linear-response formalism, we are only interested the usual transition dipole moments that are observed in one-dimensional x-ray spectroscopy. Non-orthogonality between excited state determinants and the ground state can have severely detrimental effects on transition moments [460], but a simple fix is to subtract the overlap-weighted ground-state dipole moment from the transition dipole,

$$\vec{\mu} = \sum_a X^a_i \big(\langle \Phi_0 | \hat{\mu} | \Psi^a_i \rangle - \langle \Phi_0 | \hat{\mu} | \Phi_0 \rangle \langle \Phi_0 | \Psi^a_i \rangle \big) \,, \tag{5.22}$$

where X^a_i are eigenvalues of the Tamm-Dancoff approximated Hermitian eigenvalue equation,

$$\mathbf{AX} = \omega \mathbf{X} \,. \tag{5.23}$$

This is equivalent to translating the center of charge of the molecule to the origin prior to calculating the transition moments.

5.3 Computational Details

All calculations were performed with a development version of the Q-Chem 5.4 software package [263]. The DFT calculations use a dense quadrature with 99 radial and 590 angular grid points to evaluate the exchange-correlation potential, and scalar relativistic effects are accounted for using a spin-free exact two-component (X2C) model [397]. Core-ionized references were optimized using restricted open-shell orbitals and the solutions were stabilized using combinations of maximum overlap method (MOM) [257, 315], square-gradient minimization (SGM) [318], and state-targeted energy projection (STEP) [319]. We use experimental molecular geometries whenever possible and all geometries are available in the Supporting Information.

5.4 Results and Discussion

We assess the functional dependence of EA-TDA across 15 density functionals, and while not comprehensive we include data from local density approximation (LDA), generalized gradient approximation (GGA), meta-GGA, hybrid GGA, and hybrid meta-GGA functionals. The bottom panel of Fig. 5.2 indicates a clean recovery of the Jacob's Ladder concept in DFT, with errors decreasing with each step up through the rungs. Signed errors (Fig. 5.3) reveal that semi-local functionals tend to underestimate excitation energies, whereas asymptotically correct functionals exhibit very little systematic error. Increasing the fraction of exact exchange improves error statistics up to a point with root mean squared error (RMSE) decreasing from BLYP to B3LYP and on to B5050LYP, but too much exact exchange degrades the results leading to higher RMSE for HFLYP than B5050LYP. Overall, asymptotically correct functionals perform best, and among them rCAM-B3LYP performs best of all with an RMSE of only 0.5 eV.

To understand the scope of the improvements offered by EA-TDA, we compare RMSEs of standard TDA with EA-TDA across functionals. The top panel of Fig. 5.2 reveals that for a given functional the average improvement offered by EA-TDA is on the order of tens of eV. In fact, the RMSE of rCAM-B3LYP improves by roughly two orders of magnitude, from 28.4 eV to 0.5 eV. This massive improvement is apparent in all but the SRC1 functional, which was parameterized specifically to cancel eh-SIE in standard TDDFT. This is a testament to the parameterization of SRC1, but the lack of improvement (or deterioration) of the results on switching to EA-TDA also exemplifies that eh-SIE is adequately quenched in EA-TDA.

On examination of Eq. 5.12 it becomes clear that a necessary criterion for a functional to perform well with EA-TDA is an accurate estimate of the electron affinity for each virtual MO. This is because the dominant term is $F_{ab}^{(n-1)}$, while the last two terms offer only small corrections to this energy because they are dependent on the overlap of the core orbital with the virtual MOs. Because asymptotically correct functionals perform best in the prediction of electron affinities [461, 462], so too do they perform best with EA-TDA.

The success of EA-TDA, a theory that takes orbital relaxation and eh-SIE into account,

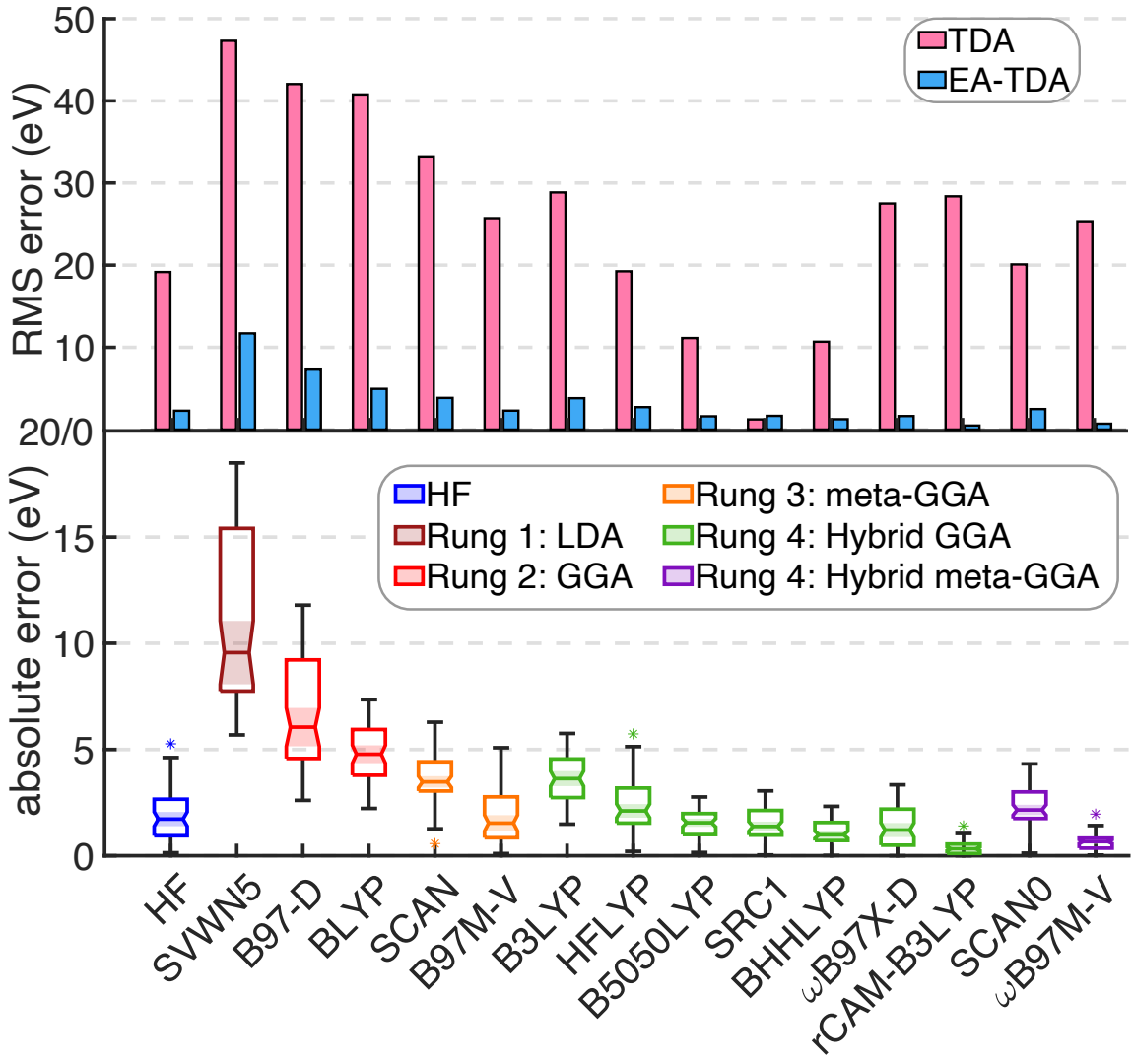


Figure 5.2: Absolute error statistics for 65 experimental K-edge transitions (lowest energy transition only). (Top) RMSE for standard TDA versus EA-TDA by functional. (Bottom) EA-TDA absolute error statistics. Upper and lower delimiters indicate maximum and minimum errors, respectively. Upper and lower bounds of each box are the upper and lower quartiles, respectively. Median absolute errors are indicated by horizontal lines and overlapping notches identify statistical similarities between distributions to the 95% confidence level. Outliers are indicated by asterisks. All calculations use aug-pcseg-1 for H and Br atoms and aug-pcX-2 otherwise.

allows us to diagnose the origins of the errors in standard TDA. Using the metric defined in 5.21 we have immediate access to the amount of eh-SIE in standard TDA, while the remainder of the error in TDA can be ascribed to orbital relaxation effects. In Fig. 5.4, we

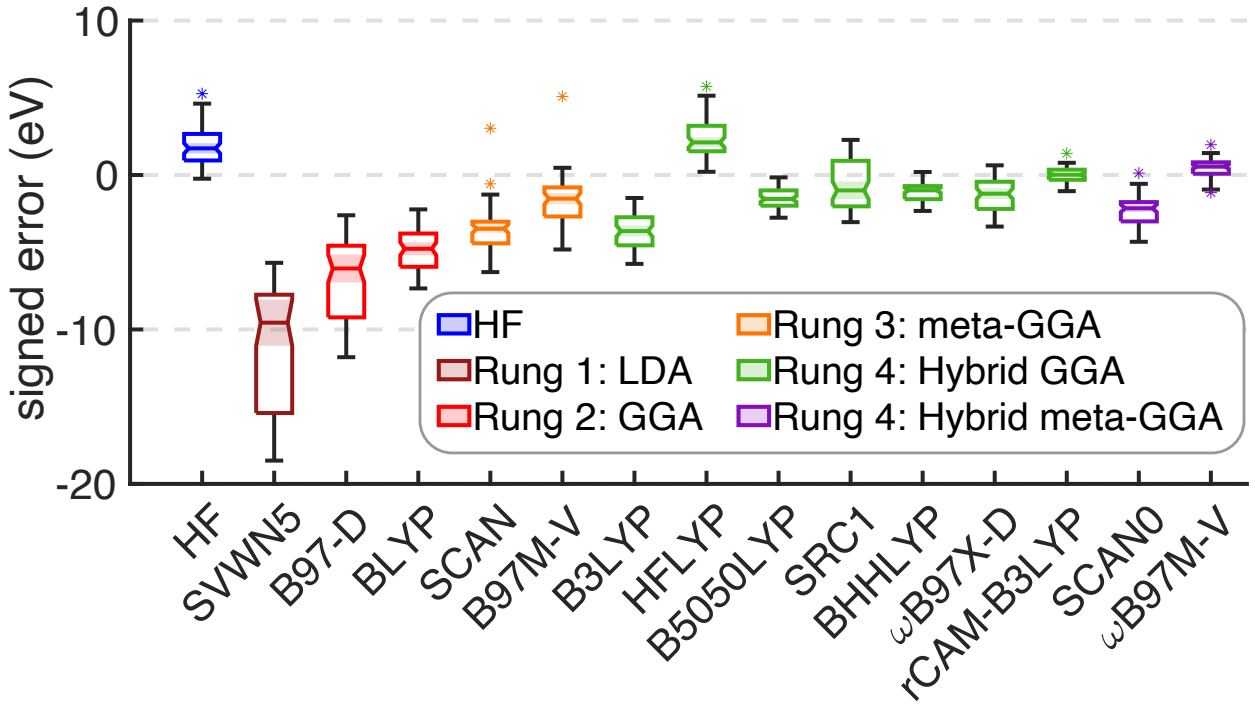


Figure 5.3: EA-TDA signed error statistics for 65 experimental K-edge transitions (lowest energy transition only). The aug-pcseg-1 basis was used for H and Br, aug-pcX-2 for all other atoms. A negative sign indicates an underestimation in the excitation energy. Upper and lower delimiters indicate maximum and minimum errors, respectively. Upper and lower bounds of each box are the upper and lower quartiles, respectively. Median absolute errors are indicated by horizontal lines and overlapping notches identify statistical similarities between distributions to the 95% confidence level. Outliers are indicated by asterisks.

define the total error (relaxation error plus eh-SIE) as the difference between standard TDA and EA-TDA, $Err(\text{TDA}) - Err(\text{EA-TDA})$, where the errors in excitation energies ω_{X} of a given method X are taken to be $Err(\text{X}) = \omega_{\text{X}} - \omega_{\text{ref}}$ with ω_{ref} representing the experimental value. Orbital relaxation contributes positive errors because without relaxation effects the predicted excitation energies are higher, whereas eh-SIE over-stabilizes the excitation energy due to a lack of particle-hole attraction, so its contribution is net negative.

The HF functional has zero contribution from eh-SIE because exact exchange yields correct particle-hole attraction. Similarly, HFLYP has a near-zero contribution from eh-SIE due to exact exchange, but it does not substantially improve upon HF in terms of orbital relaxation. Otherwise, the lion's share of error in most functionals comes from eh-SIE, with a relatively consistent contribution from a lack of orbital relaxation. An interesting exception to this rule is the SRC1 functional, which (owing to its parameterization for XAS) hosts a nearly optimal degree of error cancellation between orbital relaxation error and eh-SIE at a

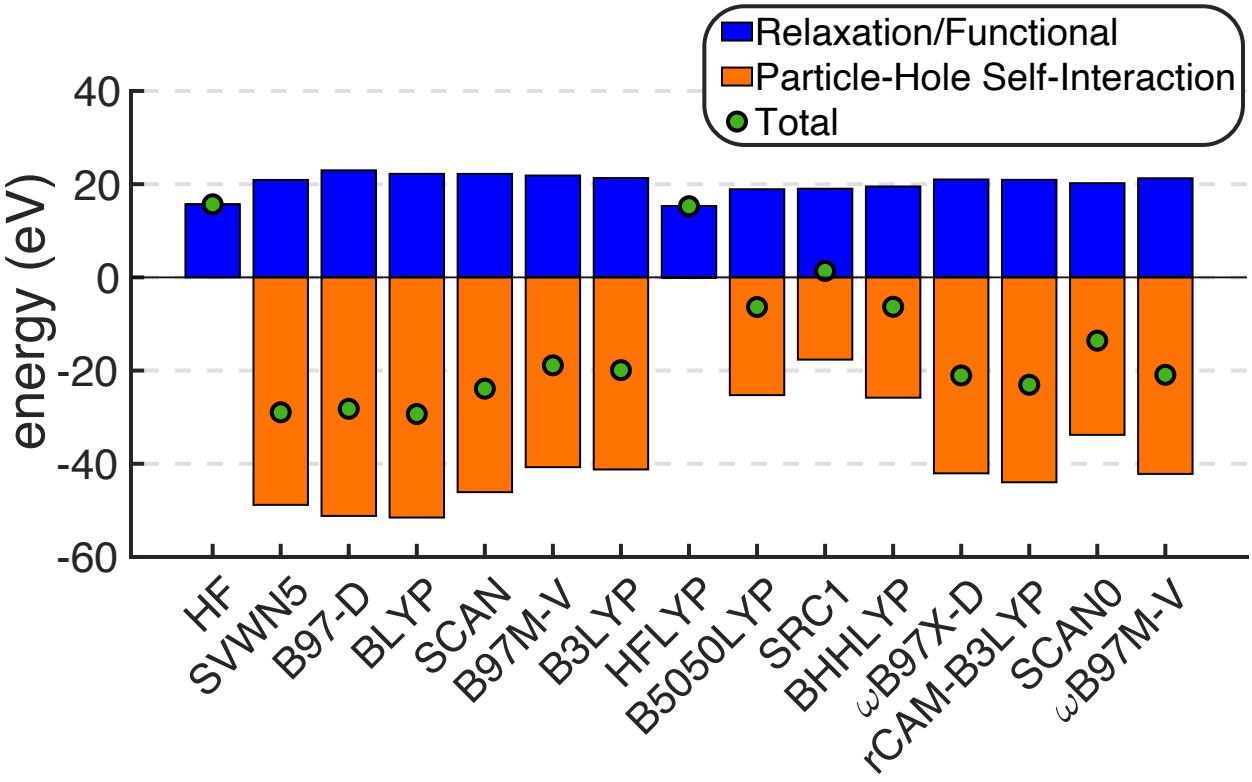


Figure 5.4: Error contributions to standard TDA that are corrected in EA-TDA, averaged over 65 K-edge transitions. The total (signed) error is taken to be the difference Err(TDA)$-$ Err(EA-TDA), which is equivalent to the sum of the bars.

respective ratio of 52 : 48. Other functionals that benefit from near-cancellation of errors are those with a large fraction of global HF exchange such as B5050LYP and BHHLYP, which explains their notably better performance in comparison to other functionals in the top panel of Fig. 5.2. Overall, the largest contribution to the errors in TDA are from eh-SIE while orbital relaxation plays a consistent, strong, but auxiliary role in defining the total error.

We repeated the statistical analysis in Fig. 5.2 for IO-TDA and standard TDA to find the best functionals for each method. Our group has previously established SCAN as an excellent functional for core-excitations with OO-DFT, so we forego further analysis here [321, 397]. The method/functional combinations that yielded the lowest errors on this test set were subjected to the more rigorous test of 132 experimental K-edge transitions of 46 molecules, ranging from 1–5 transitions per molecule. The results for IO-TDA were omitted because the best functional for IO-TDA was HF, suggesting that DFT provides no benefit to nonstationary TDA if eh-SIE is not properly taken into account. The results in Fig. 5.5 suggest that EA-TDA (RMSE = 0.5 eV) is almost as accurate as OO-DFT (RMSE = 0.4 eV) across the board, outperforming the SRC1 functionals used with standard TDA (RMSE =

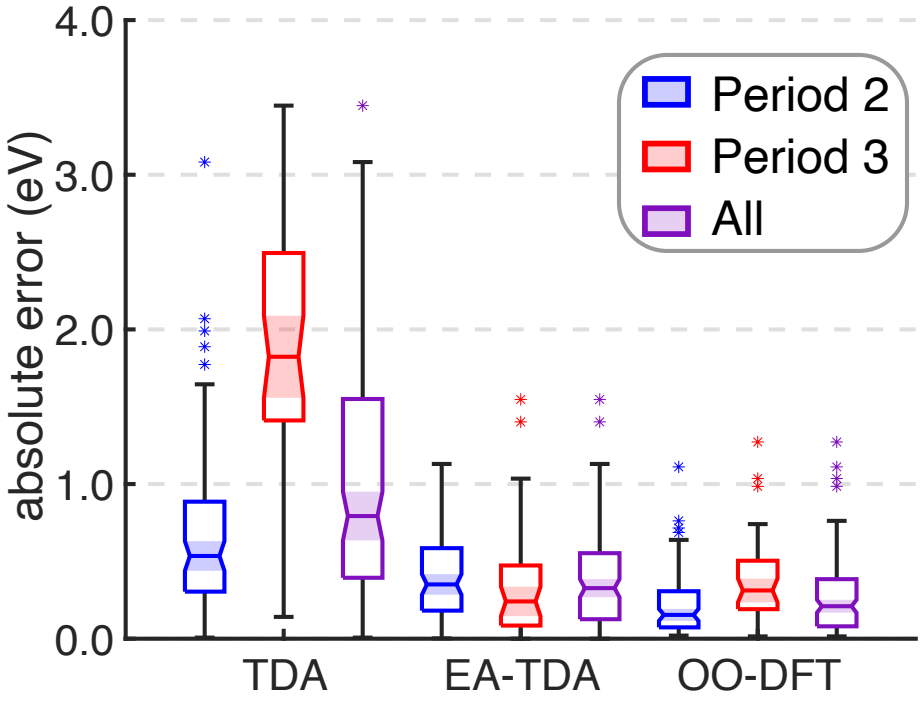


Figure 5.5: Absolute error statistics for the best method/functional combinations against 132 experimental K-edge transitions. Results are broken down by Period alongside the full data set. For TDA, SRC1-RX was used (where X = 1 or 2, depending on the period), rCAM-B3LYP and SCAN were used for EA-TDA and OO-DFT, respectively. All calculations use aug-pcseg-1 for H and Br and the doubly-augmented d-aug-pcX-2 basis otherwise.

1.3 eV) even though SRC1 was designed to accurately capture core-excitation energies.

EA-TDA also has excellent performance with respect to atomic size. While TDA and OO-DFT results show a statistically significant increase in errors from Period 2 elements to Period 3 (indicated by nonoverlapping notches in Fig. 5.5), EA-TDA results remain statistically similar with an RMSE that is equivalent to that of OO-DFT (0.5 eV). Admittedly, EA-TDA exhibits a slightly wider distribution of errors than OO-DFT, as shown by the quartiles. Being that the entire spectrum is obtained with one single EA-TDA calculation whereas OO-DFT optimizes specific configurations, this level of comparability between results is quite satisfactory.

The excellent comparability of EA-TDA and OO-DFT excitation energies begs the question: do transition properties also behave similarly? To investigate this, the experimental K-edge spectrum of ammonia is shown in Fig. 5.6 alongside calculated spectra using EA-TDA and OO-DFT. For this system, OO-DFT predicts slightly better excitation energies, whereas EA-TDA shows a slight (0.2 eV) blue shift. Insofar as transition strengths are concerned, the first major peak in the OO-DFT spectrum (normalized intensity of 0.52) corresponds

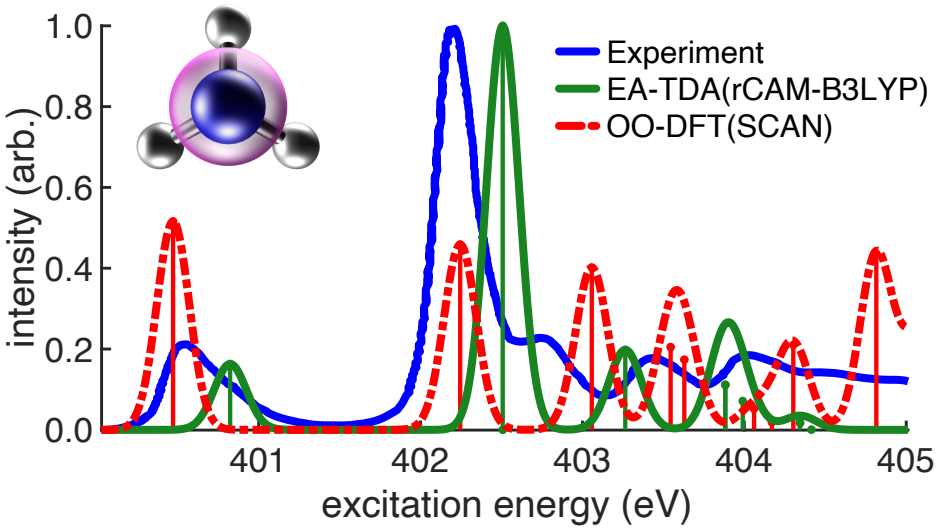


Figure 5.6: Ammonia K-edge X-ray absorption spectra for EA-TDA and OO-DFT juxtaposed against experimental data from Ref. 463. The aug-pcX-2 and aug-pcseg-1 basis sets were used for N and H, respectively.

to the 1s→3s transition, which is largely symmetry forbidden, leading to a small pre-edge peak in the experiment. The amplification of this feature is an artifact of ROKS, which allows strong singlet-triplet mixing when the initial and final states are isosymmetric [344, 464–467]. Interestingly, the errors do not seem to stem from nonorthogonality of the ROKS states, as we account for residual nonorthogonality by subtracting the overlap-weighted nuclear dipole moment [460]. Instead, these errors may emerge as a direct consequence of spurious spatial symmetry-breaking due to use of approximate functionals [468–470]. Spuriously large oscillator strengths occur frequently for 1s→3s transitions in ROKS spectra, exerting a catastrophic impact on spectra containing truly dark 1s→3s transitions like the one in *trans*-butadiene (Fig. 5.7). The EA-TDA spectrum is devoid of such spurious high-intensity dark states and the overall qualitative nature of the spectrum is captured to high fidelity in all cases.

5.5 Conclusions

Overall, we find that EA-TDA successfully ameliorates eh-SIE, providing sizable improvements over TDDFT-based response theory, yielding core-excitation energies on par with OO-DFT while avoiding the spurious high-intensity dark states that occur in the latter. We anticipate that EA-TDA, with its low computational cost and high overall accuracy, will be a tool of great importance in condensed-phase systems. Our group is currently assessing the possibility of applying EA-TDA to liquid-phase XAS [459].

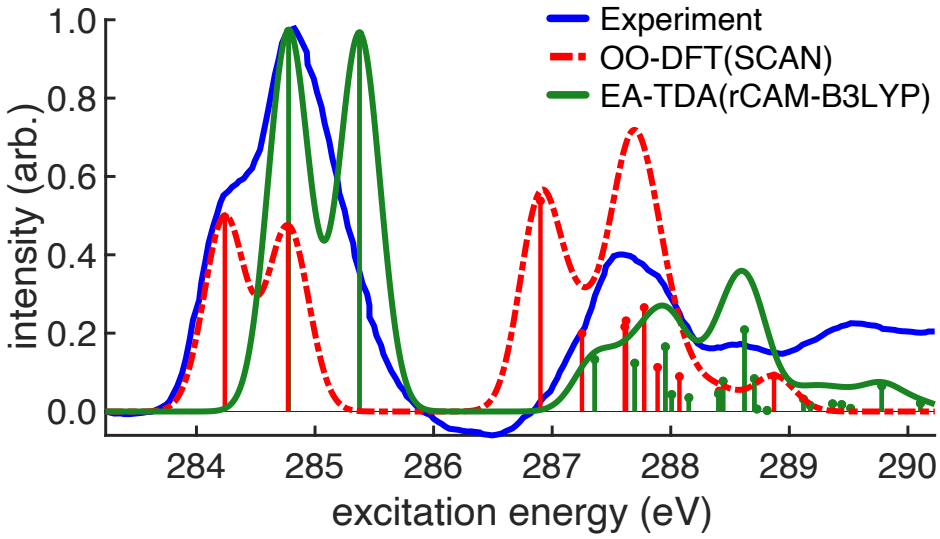


Figure 5.7: *trans*-butadiene K-edge X-ray absorption spectra for EA-TDA and OO-DFT juxtaposed against experimental data from Ref. 471. The third major peak (intensity of 0.6) in the OO-DFT spectrum corresponds to an optically dark 1s→3s transition and is absent in the EA-TDA spectrum. The aug-pcX-2 and aug-pcseg-1 basis sets were used for C and H, respectively.

Supplementary Material

Supplementary information for this work can be accessed via Ref. 472. They include additional derivations and plots, raw data and geometries of all species investigated in this work.

Chapter 6

Afterword

One never notices what has been done; one can only see what remains to be done.
Marie Curie

6.1 Overview of Published Works

This dissertation has attempted to discuss some of the modern aspects of electronic structure theory. We have presented a brief foray into the main tools we have at our disposal to solve both the non-relativistic electronic Hamiltonian, highlighting their advantages and limitations. In Chapter 2, we shed light into the symmetry-breaking breaking problem in HF theory from a fresh perspective. While it is usually discussed in the context of bond stretching and dissociation events, we proposed that there is an analog to it when static electric fields are applied to an atom, and that the choice of HF orbitals matters when trying to describe strong-field ionization processes. Acknowledging the limitation of our "race-track" of basis functions approach, we emphasize the need for better hybrid Gaussian-plane waves basis sets and appropriate orbitals to model processes, both in the static and time-dependent domains, involving ultra-strong electric fields.

Starting in Chapter 3, we turned our attention to analyzing the important aspects of computational modeling of core-spectroscopy. There, we investigated the key ingredients and approximations necessary to properly capture oscillator strengths and intensities in the context of X-ray emission spectroscopy (XES). While the state-specific route offered by OODFT seems appealing, the fact that each state has to be targeted individually complicates matters. On the other hand, TDDFT is insufficient for certain systems. We tackled these two issues by proposing a powerful, yet simple, metric that indicates where TDDFT might fail and by developing the MBXES approach. In MBXES, only two state-specific calculations are required (the ground-state and the core-ion state), and the intensities of the XES spectrum can be calculated using these two sets of orbitals.

In XES we can recover information about the occupied energy levels of a system. A step

further would be to investigate the electronic structure of the unoccupied levels, which can be accomplished in X-ray absorption spectroscopy (XAS). This is the main motivation for the tools we have developed in Chapters 4 and 5. In Chapter 4, we extended the applicability of OODFT to heavier elements by incorporating scalar-relativistic effects through X2C. Despite success in accurately describing excitation energies and spectral features of systems containing elements of the second and third rows of the periodic table, X2C-OODFT starts to run out of steam for the first row of transition metals. In Chapter 5, we proposed a new framework within the linear-response approach to tackle XAS. Inspired by the findings in the previous chapter, we devised a double response formalism that effectively creates particles in the context of TDDFT. This method, dubbed EA-TDDFT, is as accurate as OODFT for excitation energies of core-excited states, while having the advantage of providing all of the excited states and the full spectrum in a single calculation. Instead of competing approaches, OODFT and EA-TDDFT should be regarded as complementary tools for computational XAS. They showcase the fertile ground that new developments in spectroscopy provides to theoretical chemists. We emphasize that such synergy is of paramount importance to advance science as a whole.

Finally, even though the previous chapters offer a summary of the main thrusts explored in dissertation, we would like to highlight a few additional published works that encompass important developments in modern quantum chemistry:

1. Arias-Martinez, J. E.; **Cunha, L. A.**; Oosterbaan, K. J.; Lee, J.; Head-Gordon, M. "Accurate core excitation and ionization energies from a state-specific coupled-cluster singles and doubles approach" *Phys. Chem. Chem. Phys.*, **24**, 10728–20741, 2022.

   So far, we have explored computational core-spectroscopy through the lens of DFT. However, as noted in Chapter 1, wavefunction formalisms are equally important to modeling excited-states in general. Hence, we have developed a state-specific approach based on coupled-cluster theory to model XAS.

2. Lee, J; Feng, X.; **Cunha, L. A.**; Gonthier, J. F.; Epifanovsky, E.; Head-Gordon, M. "Approaching the basis set limit in Gaussian-orbital-based periodic calculations with transferability: Performance of pure density functionals for simple semiconductors" *J. Chem. Phys.*, **155**, 164102, 2021.

   This dissertation has only discussed aspects of the electronic structure of molecular systems. While the condensed matter community has relied largely on simulations using plane-waves, localized basis sets can offer certain advantages going beyond the mean-field approach. Hence, we started developing a code that uses Gaussian-type orbitals for extended systems, first analyzing the behavior of semiconductor gaps with modern local and semi-local DFAs.

3. Rossomme, E.; **Cunha, L. A.**; Li, W.; Chen, K.; McIsaac, A. R.; Head-Gordon, T.; Head-Gordon, M. "The Good, the Bad, and the Ugly: Pseudopotential Inconsistency

Errors in Molecular Applications of Density Functional Theory" *J. Chem. Theory Comput.*, **19**, 10, 2827-2841, 2023.

In Chapter 1, we discussed what we understand as the three pillars of modern quantum chemistry. This dissertation has presented developments along the correlation and the relativistic axes of Fig 1.3. However, designing better approximations to deal with the basis set axis is necessary in an attempt to lower the computational cost of electronic structure methods. Among them, pseudopotentials that replace certain electrons by an effective potential in the Hamiltonian are widely used tools both in the molecular and condensed matter communities. We have analyzed the accuracy and transferability of pseudopotentials for different chemical processes.

6.2 Remaining Challenges

Our starting point in this journey was the Schrödinger equation and the Hamiltonian that described the behavior of both nuclei and electrons. Exploring the separation of time-scales between these two sets of particles allows us to focus solely on the electronic structure problem, either in its non-relativistic or relativistic form. We have discussed tools that are capable of providing accurate energies and properties for ground-states and excited states involved in a plethora of chemical processes, expanding our computational toolbox and allowing us to comprehend nature all around us. Have we finally reached the point where Dirac's quote that opened this dissertation in Chapter 1 is no longer valid?

Once again, the perspectives of an optimistic aspiring quantum chemist prevail: there is a lot more to explore! New theoretical models spring from novel experiments and vice-versa, setting up the perfect cycle of discovery and innovation. It would be foolish to attempt to provide an extensive list of the challenges yet to be solved in electronic structure or quantum chemistry in general. Such a list is, fortunately, far too long already and, likely, will never be comprehensive. We, however, highlight a few topics that we believe will receive growing interest in the coming years:

1. *Opening the quantum chemical toolbox to extended systems*

   As indicated previously, there are certain advantages of using localized basis sets in quantum chemistry, even for systems that have been historically employed plane-wave simulations, such as periodic systems. For instance, it is known that correlation effects converge faster towards the complete basis limit when a localized basis is employed [473]. Moreover, condensed matter physicists have been fascinated by the emergence of new materials whose properties arise as a consequence, to some extent, of correlation effects. For instance, it is argued that the mechanism of high temperature superconductivity in cuprates requires considering strong correlation effects [474], which are beyond of the mean-field treatment in DFT, the workhorse for periodic systems. Quantum chemists have at their disposal a somewhat extensive toolbox to treat

strong correlation [39, 40, 42, 50] (although there is always room for improvements and new tools), hence we highlight that extending this machinery to extended system could help interpret some of the arising new phenomena.

2. *Expanding the Hamiltonian and multiscale modeling*

   Quantum chemists have long been interested in the behavior of electrons. Some have dared to go beyond and also solved the nuclear Schrödinger equation, but few have attempted to incorporate nuclear and electronic effects on equal footing [475, 476]. Even fewer have attempted to explore the consequences of including even more particles, such as photons, in their Hamiltonian [477–479]. However, understanding the behavior of electrons when coupled to nuclei and/or photons is of paramount importance to unravel new ways of controlling matter [480]. For instance, some mechanisms of superconductivity are mediated through phonons, or lattice (nuclei) vibrations [481, 482]. We believe this sort of multiscale simulation, where electrons, nuclei and photons are treated on equal footing, is one of the frontiers of quantum chemical development.

3. *Combining different tools through embedding approaches*

   Finally, it is not enough to simply devise new computational tools. They need to be applied to understand the properties of large systems and complex chemical and physical processes. Including more physics in a model should not be a barrier towards accomplishing this task. It is necessary to include these important physical effects where they matter the most. For instance, it is argued that relativistic effects are rather local, so schemes that explore this consequence have been proposed to efficiently model spin-orbit coupling [130]. We believe embedding techniques [49, 483–485] could play an important role in helping us finally solve the complex equations of molecular quantum mechanics and understand nature around us.

Bibliography

(1) Ehrenfest, P. Welche Züge der Lichtquantenhypothese spielen in der Theorie der Wärmestrahlung eine wesentliche Rolle? *Ann. Phys. (Berl.)* **1911**, *341*, 91–118.

(2) Planck, M. Ueber das Gesetz der Energieuerteilzcng im Norrnalspectrzcm. *Ann. Phys. (Berl.)* **1901**, *309*, 553–563.

(3) Feynman, R.; Leighton, R.; Sands, M., *The Feynman Lectures on Physics*; The Feynman Lectures on Physics v. 1; Addison-Wesley Publishing Company: 1965.

(4) Bohr, N. I. On the constitution of atoms and molecules. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **1913**, *26*, 1–25.

(5) Griffiths, D. J.; Schroeter, D. F., *Introduction to Quantum Mechanics*, 3rd ed.; Cambridge University Press: 2018.

(6) Piela, L., *Ideas of quantum chemistry*; Elsevier: 2006.

(7) Einstein, A. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Ann. Phys. (Berl.)* **1905**, *4*.

(8) Taylor, J. R.; Taylor, J. R., *Classical mechanics*; University Science Books: 2005; Vol. 1.

(9) Born, M.; Oppenheimer, R. Zur quantentheorie der molekeln. *Ann. Phys. (Berl.)* **1927**, *389*, 457–484.

(10) Tully, J. C. Perspective on "zur quantentheorie der molekeln". *Theor. Chem. Acc.* **2000**, *103*, 173–176.

(11) Yarkony, D. R. Diabolical conical intersections. *Rev. Mod. Phys.* **1996**, *68*, 985.

(12) Yarkony, D. R. Conical intersections: The new conventional wisdom. *J. Phys. Chem. A* **2001**, *105*, 6277–6293.

(13) Worth, G. A.; Cederbaum, L. S. Beyond Born-Oppenheimer: molecular dynamics through a conical intersection. *Annu. Rev. Phys. Chem.* **2004**, *55*, 127–158.

(14) Fermi, E. Sulla quantizzazione del gas perfetto monoatomico. *Rend. Lincei* **1926**, *145*.

(15) Dirac, P. A. M. On the theory of quantum mechanics. *Proc. R. Soc. A* **1926**, *112*, 661–677.

(16) Pauli, W. Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren. *Z. Phys.* **1925**, *31*, 765–783.

(17) Slater, J. C. The theory of complex spectra. *Phys. Rev.* **1929**, *34*, 1293.

(18) Hammes-Schiffer, S.; Andersen, H. C. The advantages of the general Hartree–Fock method for future computer simulation of materials. *J. Chem. Phys.* **1993**, *99*, 1901–1913.

(19) Henderson, T. M.; Jimenez-Hoyos, C. A.; Scuseria, G. E. Magnetic structure of density matrices. *J. Chem. Theory Comput.* **2018**, *14*, 649–659.

(20) Hartree, D. R. In *Math. Proc. Camb. Philos. Soc.* 1928; Vol. 24, pp 89–110.

(21) Fock, V. Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems. *Z. Phys.* **1930**, *61*, 126–148.

(22) Fock, V. „Selfconsistent field "mit Austausch für Natrium. *Z. Phys.* **1930**, *62*, 795–805.

(23) Slater, J. C. Note on Hartree's method. *Phys. Rev.* **1930**, *35*, 210.

(24) Szabo, A.; Ostlund, N. S., *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover Publications, Inc.: Mineola, New York, 1996, pp 286–296.

(25) Helgaker, T.; Jorgensen, P.; Olsen, J., *Molecular electronic-structure theory*; John Wiley & Sons: 2013.

(26) Koopmans, T. Über die Zuordnung von Wellenfunktionen und Eigenwerten zu den einzelnen Elektronen eines Atoms. *Physica* **1934**, *1*, 104–113.

(27) Löwdin, P.-O. Quantum theory of many-particle systems. III. Extension of the Hartree-Fock scheme to include degenerate systems and correlation effects. *Phys. Rev.* **1955**, *97*, 1509.

(28) Schrödinger, E. Quantisierung als eigenwertproblem. *Ann. Phys. (Berl.)* **1926**, *385*, 437–490.

(29) Møller, C.; Plesset, M. S. Note on an approximation treatment for many-electron systems. *Phys. Rev.* **1934**, *46*, 618.

(30) Cremer, D. Møller–Plesset perturbation theory: from small molecule methods to methods for thousands of atoms. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 509–530.

(31) Lennard-Jones, J. E. Perturbation problems in quantum mechanics. *Proc. R. Soc. Lond. A* **1930**, *129*, 598–615.

(32) Brillouin, L. Les problèmes de perturbations et les champs self-consistents. *J. Phys. Radium* **1932**, *3*, 373–389.

(33) Wigner, E. In *The Collected Works of Eugene Paul Wigner*, ed. by Wightman, A. S., Springer: 1997, pp 131–136.

(34) Wilson, S.; Hubač, I.; Mach, P.; Pittner, J.; Čársky, P. In *Advanced Topics in Theoretical Chemical Physics*, 2003, pp 71–117.

(35) Carter-Fenk, K.; Head-Gordon, M. Repartitioned Brillouin-Wigner perturbation theory with a size-consistent second-order correlation energy. *J. Chem. Phys.* **2023**, *158*.

(36) Kurlancheek, W.; Head-Gordon, M. Violations of N-representability from spin-unrestricted orbitals in Møller–Plesset perturbation theory and related double-hybrid density functional theory. *Mol. Phys.* **2009**, *107*, 1223–1232.

(37) Shavitt, I. In *Methods of electronic structure theory*; Springer: 1977, pp 189–275.

(38) Sherrill, C. D.; Schaefer III, H. F. In *Advances in quantum chemistry*; Elsevier: 1999; Vol. 34, pp 143–269.

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

(40) Evangelista, F. A. Adaptive multiconfigurational wave functions. *J. Chem. Phys.* **2014**, *140*.

(41) Holmes, A. A.; Tubman, N. M.; Umrigar, C. Heat-bath configuration interaction: An efficient selected configuration interaction algorithm inspired by heat-bath sampling. *J. Chem. Theory Comput.* **2016**, *12*, 3674–3680.

(42) Tubman, N. M.; Freeman, C. D.; Levine, D. S.; Hait, D.; Head-Gordon, M.; Whaley, K. B. Modern approaches to exact diagonalization and selected configuration interaction with the adaptive sampling CI method. *J. Chem. Theory Comput.* **2020**, *16*, 2139–2159.

(43) Čížek, J. On the correlation problem in atomic and molecular systems. Calculation of wavefunction components in Ursell-type expansion using quantum-field theoretical methods. *J. Chem. Phys.* **1966**, *45*, 4256–4266.

(44) Bartlett, R. J. Many-body perturbation theory and coupled cluster theory for electron correlation in molecules. *Annu. Rev. Phys. Chem.* **1981**, *32*, 359–401.

(45) Bartlett, R. J.; Musiał, M. Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.* **2007**, *79*, 291.

(46) Shavitt, I.; Bartlett, R. J., *Many-body methods in chemistry and physics: MBPT and coupled-cluster theory*; Cambridge university press: 2009.

(47) Bartlett, R. J.; Noga, J. The expectation value coupled-cluster method and analytical energy derivatives. *Chem. Phys. Lett.* **1988**, *150*, 29–36.

(48) Szalay, P. G.; Nooijen, M.; Bartlett, R. J. Alternative ansätze in single reference coupled-cluster theory. III. A critical analysis of different methods. *J. Chem. Phys.* **1995**, *103*, 281–298.

(49) Lin, N.; Marianetti, C.; Millis, A. J.; Reichman, D. R. Dynamical mean-field theory for quantum chemistry. *Physical review letters* **2011**, *106*, 096402.

(50) Freitag, L.; Reiher, M. In *Quantum Chemistry and Dynamics of Excited States*; John Wiley & Sons, Ltd: 2020; Chapter 7, pp 205–245.

(51) Thomas, L. H. In *Math. Proc. Camb. Philos. Soc.* 1927; Vol. 23, pp 542–548.

(52) Fermi, E. Un metodo statistico per la determinazione di alcune priorieta dell'atome. *Rend. Accad. Naz. Lincei* **1927**, *6*, 32.

(53) Dirac, P. A. M. In *Math. Proc. Camb. Philos. Soc.* 1930; Vol. 26, pp 376–385.

(54) Teller, E. On the stability of molecules in the Thomas-Fermi theory. *Rev. Mod. Phys.* **1962**, *34*, 627.

(55) Hohenberg, P.; Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **1964**, *136*, B864.

(56) Levy, M. Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 6062–6065.

(57) Weizsäcker, C. v. Zur theorie der kernmassen. *Z. Phys.* **1935**, *96*, 431–458.

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

(59) Kohn, W. Nobel Lecture: Electronic structure of matter—wave functions and density functionals. *Rev. Mod. Phys.* **1999**, *71*, 1253.

(60) Mardirossian, N.; Head-Gordon, M. Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals. *Mol. Phys.* **2017**, *115*, 2315–2372.

(61) Lehtola, S.; Steigemann, C.; Oliveira, M. J.; Marques, M. A. Recent developments in libxc—A comprehensive library of functionals for density functional theory. *SoftwareX* **2018**, *7*, 1–5.

(62) Perdew, J. P.; Schmidt, K. In *AIP Conference Proceedings*, 2001; Vol. 577, pp 1–20.

(63) Pines, D. A collective description of electron interactions: IV. Electron interaction in metals. *Phys. Rev.* **1953**, *92*, 626.

(64) Ceperley, D. M.; Alder, B. J. Ground state of the electron gas by a stochastic method. *Phys. Rev. Lett.* **1980**, *45*, 566.

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

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

(67) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865.

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

(69) Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.* **2006**, *27*, 1787–1799.

(70) Perdew, J. P.; Constantin, L. A. Laplacian-level density functionals for the kinetic energy density and exchange-correlation energy. *Phys. Rev. B* **2007**, *75*, 155109.

(71) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the density functional ladder: Nonempirical meta–generalized gradient approximation designed for molecules and solids. *Phys. Rev. Lett.* **2003**, *91*, 146401.

(72) Sun, J.; Ruzsinszky, A.; Perdew, J. P. Strongly Constrained and Appropriately Normed Semilocal Density Functional. *Phys. Rev. Lett.* **2015**, *115*, 036402.

(73) Mardirossian, N.; Head-Gordon, M. Mapping the genome of meta-generalized gradient approximation density functionals: The search for B97M-V. *J. Chem. Phys.* **2015**, *142*, 074111.

(74) Perdew, J. P.; Parr, R. G.; Levy, M.; Balduz Jr, J. L. Density-functional theory for fractional particle number: derivative discontinuities of the energy. *Phys. Rev. Lett.* **1982**, *49*, 1691.

(75) Iikura, H.; Tsuneda, T.; Yanai, T.; Hirao, K. A long-range correction scheme for generalized-gradient-approximation exchange functionals. *J. Chem. Phys.* **2001**, *115*, 3540–3544.

(76) Tawada, Y.; Tsuneda, T.; Yanagisawa, S.; Yanai, T.; Hirao, K. A long-range-corrected time-dependent density functional theory. *J. Chem. Phys.* **2004**, *120*, 8425–8433.

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

(78) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158–6170.

(79) Mardirossian, N.; Head-Gordon, M. ω B97M-V: A combinatorially optimized, range-separated hybrid, meta-GGA density functional with VV10 nonlocal correlation. *J. Chem. Phys.* **2016**, *144*, 214110.

(80) Jaramillo, J.; Scuseria, G. E.; Ernzerhof, M. Local hybrid functionals. *J. Chem. Phys.* **2003**, *118*, 1068–1073.

(81) Maier, T. M.; Arbuznikov, A. V.; Kaupp, M. Local hybrid functionals: Theory, implementation, and performance of an emerging new tool in quantum chemistry and beyond. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2019**, *9*, e1378.

(82) Haasler, M.; Maier, T. M.; Grotjahn, R.; Gückel, S.; Arbuznikov, A. V.; Kaupp, M. A local hybrid functional with wide applicability and good balance between (de) localization and left–right correlation. *J. Chem. Theory Comput.* **2020**, *16*, 5645–5657.

(83) Grimme, S. Semiempirical hybrid density functional with perturbative second-order correlation. *J. Chem. Phys.* **2006**, *124*, 034108.

(84) Mardirossian, N.; Head-Gordon, M. Survival of the most transferable at the top of Jacob's ladder: Defining and testing the ω B97M (2) double hybrid density functional. *J. Chem. Phys.* **2018**, *148*, 241736.

(85) Orenstein, J. Ultrafast spectroscopy of quantum materials. *Phys. Today* **2012**, *65*, 44–50.

(86) Maiuri, M.; Garavelli, M.; Cerullo, G. Ultrafast spectroscopy: State of the art and open challenges. *J. Am. Chem. Soc.* **2019**, *142*, 3–15.

(87) Kim, J. Y.; Lee, J.-W.; Jung, H. S.; Shin, H.; Park, N.-G. High-efficiency perovskite solar cells. *Chem. Rev.* **2020**, *120*, 7867–7918.

(88) Dreuw, A.; Wormit, M. The algebraic diagrammatic construction scheme for the polarization propagator for the calculation of excited states. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2015**, *5*, 82–95.

(89) Krylov, A. I. Equation-of-motion coupled-cluster methods for open-shell and electronically excited species: The hitchhiker's guide to Fock space. *Annu. Rev. Phys. Chem.* **2008**, *59*, 433–462.

(90) Bartlett, R. J. Coupled-cluster theory and its equation-of-motion extensions. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 126–138.

(91) Dreuw, A.; Head-Gordon, M. Single-reference ab initio methods for the calculation of excited states of large molecules. *Chem. Rev.* **2005**, *105*, 4009–4037.

(92) Casida, M. E.; Huix-Rotllant, M. Progress in time-dependent density-functional theory. *Annu. Rev. Phys. Chem.* **2012**, *63*, 287–323.

(93) Clune, R.; Shea, J. A.; Neuscamman, E. N5-scaling excited-state-specific perturbation theory. *J. Chem. Theory. Comput.* **2020**, *16*, 6132–6141.

(94) Lee, J.; Small, D. W.; Head-Gordon, M. Excited states via coupled cluster theory without equation-of-motion methods: Seeking higher roots with application to doubly excited states and double core hole states. *J. Chem. Phys.* **2019**, *151*.

(95) Arias-Martinez, J. E.; Cunha, L. A.; Oosterbaan, K. J.; Lee, J.; Head-Gordon, M. Accurate core excitation and ionization energies from a state-specific coupled-cluster singles and doubles approach. *Phys. Chem. Chem. Phys.* **2022**, *24*, 20728–20741.

(96) Matuszek, A. M.; Reynisson, J. Defining known drug space using DFT. *Mol. Inform.* **2016**, *35*, 46–53.

(97) Pyzer-Knapp, E. O.; Suh, C.; Gómez-Bombarelli, R.; Aguilera-Iparraguirre, J.; Aspuru-Guzik, A. What is high-throughput virtual screening? A perspective from organic materials discovery. *Annu. Rev. Mat. Res.* **2015**, *45*, 195–216.

(98) Hafiz, H. et al. A high-throughput data analysis and materials discovery tool for strongly correlated materials. *npj Comp. Mat.* **2018**, *4*, 63.

(99) Pyykko, P. Relativistic effects in structural chemistry. *Chem. Rev.* **1988**, *88*, 563–594.

(100) Pyykkö, P. Relativistic effects in chemistry: more common than you thought. *Annu. Rev. Phys. Chem.* **2012**, *63*, 45–64.

(101) Löffelsender, S.; Schwerdtfeger, P.; Grimme, S.; Mewes, J.-M. It's Complicated: On Relativistic Effects and Periodic Trends in the Melting and Boiling Points of the Group 11 Coinage Metals. *J. Am. Chem. Soc.* **2021**, *144*, 485–494.

(102) Gerlach, W.; Stern, O. Der experimentelle nachweis der richtungsquantelung im magnetfeld. *Z. Phys.* **1922**, *9*, 349–352.

(103) Pauli Jr, W. 1927,'Zur Quantenmechanik des magnetischen Elektrons'. *Z. Phys* **1927**, *43*, 601.

(104) Greiner, W. et al., *Relativistic quantum mechanics*; Springer: 2000; Vol. 2.

(105) Dyall, K. G.; Fægri Jr, K., *Introduction to relativistic quantum chemistry*; Oxford University Press: 2007.

(106) Reiher, M.; Wolf, A., *Relativistic quantum chemistry: the fundamental theory of molecular science*; John Wiley & Sons: 2014.

(107) Liu, W. Essentials of relativistic quantum chemistry. *The Journal of chemical physics* **2020**, *152*.

(108) Klein, O. Quantentheorie und fünfdimensionale Relativitätstheorie. *Z. Phys.* **1926**, *37*, 895–906.

(109) Gordon, W. Der comptoneffekt nach der schrödingerschen theorie. *Z. Phys.* **1926**, *40*, 117–133.

(110) Dirac, P. A. M. The quantum theory of the electron. *Proc. R. Soc. Lond. A* **1928**, *117*, 610–624.

(111) Foldy, L. L.; Wouthuysen, S. A. On the Dirac theory of spin 1/2 particles and its non-relativistic limit. *Phys. Rev.* **1950**, *78*, 29.

(112) Foldy, L. The electromagnetic properties of Dirac particles. *Phys. Rev.* **1952**, *87*, 688.

(113) Lamb Jr, W. E.; Retherford, R. C. Fine structure of the hydrogen atom by a microwave method. *Phys. Rev.* **1947**, *72*, 241.

(114) Bethe, H. A. The electromagnetic shift of energy levels. *Phys. Rev.* **1947**, *72*, 339.

(115) Schwinger, J. Quantum electrodynamics. I. A covariant formulation. *Phys. Rev.* **1948**, *74*, 1439.

(116) Schwinger, J. On quantum-electrodynamics and the magnetic moment of the electron. *Phys. Rev.* **1948**, *73*, 416.

(117) Tomonaga, S.-i. On a relativistically invariant formulation of the quantum theory of wave fields. *Prog. Theor. Phys.* **1946**, *1*, 27–42.

(118) Feynman, R. P. Mathematical formulation of the quantum theory of electromagnetic interaction. *Phys. Rev.* **1950**, *80*, 440.

(119) Craig, D. P.; Thirunamachandran, T., *Molecular quantum electrodynamics: an introduction to radiation-molecule interactions*; Dover Publications: 1998.

(120) Breit, G. Dirac's equation and the spin-spin interactions of two electrons. *Phys. Rev.* **1932**, *39*, 616.

(121) Liu, J.; Cheng, L. Relativistic coupled-cluster and equation-of-motion coupled-cluster methods. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2021**, *11*, e1536.

(122) Bethe, H. A.; Salpeter, E. E., *Quantum mechanics of one-and two-electron atoms*; Springer Science & Business Media: 2012.

(123) Douglas, M.; Kroll, N. M. Quantum electrodynamical corrections to the fine structure of helium. *Ann. Phys.* **1974**, *82*, 89–155.

(124) Hess, B. A. Relativistic electronic-structure calculations employing a two-component no-pair formalism with external-field projection operators. *Phys. Rev. A* **1986**, *33*, 3742.

(125) Reiher, M. Douglas–Kroll–Hess Theory: a relativistic electrons-only theory for chemistry. *Theo. Chem. Acc.* **2006**, *116*, 241–252.

(126) Nakajima, T.; Hirao, K. The Douglas–Kroll–Hess Approach. *Chem. Rev.* **2012**, *112*, 385–402.

(127) Dyall, K. G. Interfacing relativistic and nonrelativistic methods. I. Normalized elimination of the small component in the modified Dirac equation. *J. Chem. Phys.* **1997**, *106*, 9618–9626.

(128) Saue, T. Relativistic Hamiltonians for Chemistry: A Primer. *Chem. Phys. Chem.* **2011**, *12*, 3077–3094.

(129) Kutzelnigg, W. Basis set expansion of the Dirac operator without variational collapse. *Int. J. Quantum Chem.* **1984**, *25*, 107–129.

(130) Liu, J.; Cheng, L. An atomic mean-field spin-orbit approach within exact two-component theory for a non-perturbative treatment of spin-orbit coupling. *J. Chem. Phys.* **2018**, *148*.

(131) Zhang, C.; Cheng, L. Atomic mean-field approach within exact two-component theory based on the Dirac–Coulomb–Breit Hamiltonian. *J. Phys. Chem. A* **2022**, *126*, 4537–4553.

(132) Cohen, M. L.; Louie, S. G., *Fundamentals of condensed matter physics*; Cambridge University Press: 2016.

(133) Andrade, X. et al. Real-space grids and the Octopus code as tools for the development of new simulation approaches for electronic systems. *Phys. Chem. Chem. Phys.* **2015**, *17*, 31371–31396.

(134) Gill, P. M. In *Advances in quantum chemistry*; Elsevier: 1994; Vol. 25, pp 141–205.

(135) Wang, Z.; Aldossary, A.; Head-Gordon, M. Sparsity of the electron repulsion integral tensor using different localized virtual orbital representations in local second-order Møller–Plesset theory. *J. Chem. Phys.* **2023**, *158*.

(136) Lee, J.; Head-Gordon, M. Regularized orbital-optimized second-order Møller–Plesset perturbation theory: A reliable fifth-order-scaling electron correlation model with orbital energy dependent regularizers. *J. Chem. Theory Comput.* **2018**, *14*, 5203–5219.

(137) Pederson, R.; Kalita, B.; Burke, K. Machine learning and density functional theory. *Nat. Rev. Phys.* **2022**, *4*, 357–358.

(138) Eisfeld, W.; Morokuma, K. A Detailed Study on the Symmetry Breaking and its Effect on the Potential Surface of NO 3. *J. Chem. Phys.* **2000**, *113*, 5587–5597.

(139) Sherrill, C. D.; Lee, M. S.; Head-Gordon, M. On the performance of density functional theory for symmetry-breaking problems. *Chem. Phys. Lett.* **1999**, *302*, 425–430.

(140) Russ, N. J.; Crawford, T. D.; Tschumper, G. S. Real versus artifactual symmetry-breaking effects in Hartree–Fock, density-functional, and coupled-cluster methods. *J. Chem. Phys.* **2004**, *120*, 7298–7306.

(141) Dobaczewski, J.; Dudek, J.; Rohoziński, S.; Werner, T. Point symmetries in the Hartree-Fock approach. II. Symmetry-breaking schemes. *Phys. Rev. C* **2000**, *62*, 014311.

(142) Hait, D.; Rettig, A.; Head-Gordon, M. Well-behaved versus ill-behaved density functionals for single bond dissociation: Separating success from disaster functional by functional for stretched H2. *J. Chem. Phys.* **2019**, *150*, 094115.

(143) Hait, D.; Rettig, A.; Head-Gordon, M. Beyond the Coulson–Fischer point: characterizing single excitation CI and TDDFT for excited states in single bond dissociations. *Phys. Chem. Chem. Phys* **2019**, *21*, 21761–21775.

(144) Lykos, P.; Pratt, G. W. Discussion on The Hartree-Fock Approximation. *Rev. Mod. Phys.* **1963**, *35*, 496–501.

(145) Fukutome, H. Unrestricted Hartree–Fock theory and its applications to molecules and chemical reactions. *Int. J. Quantum Chem.* **1981**, *20*, 955–1065.

(146) Stuber, J. L.; Paldus, J. In *Fundamental World of Quantum Chemistry*, Brändas, E. J., Kryachko, E. S., Eds.; Kluwer Academic Publishers: 2003; Vol. 1; Chapter 4, pp 67–139.

(147) Ebata, S.; Nakatsukasa, T.; Inakura, T.; Yoshida, K.; Hashimoto, Y.; Yabana, K. Canonical-basis time-dependent Hartree-Fock-Bogoliubov theory and linear-response calculations. *Phys. Rev. C* **2010**, *82*, 034306.

(148) Bertsch, G.; Robledo, L. M. Symmetry restoration in Hartree-Fock-Bogoliubov based theories. *Phys. Rev. Lett.* **2012**, *108*, 042505.

(149) Thouless, D. Stability conditions and nuclear rotations in the Hartree-Fock theory. *Nuc. Phys.* **1960**, *21*, 225–232.

(150) Seeger, R.; Pople, J. A. Self-consistent molecular orbital methods. XVIII. Constraints and stability in Hartree–Fock theory. *J. Chem. Phys.* **1977**, *66*, 3045–3050.

(151) Coulson, P. C.; Fischer, M. I. XXXIV. Notes on the molecular orbital treatment of the hydrogen molecule. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **1949**, *40*, 386–393.

(152) Čížek, J.; Paldus, J. Stability Conditions for the Solutions of the Hartree—Fock Equations for Atomic and Molecular Systems. Application to the Pi-Electron Model of Cyclic Polyenes. *J. Chem. Phys.* **1967**, *47*, 3976–3985.

(153) Tóth, Z.; Pulay, P. Finding symmetry breaking Hartree-Fock solutions: The case of triplet instability. *J. Chem. Phys.* **2016**, *145*, 164102.

(154) Stück, D.; Baker, T. A.; Zimmerman, P.; Kurlancheek, W.; Head-Gordon, M. On the nature of electron correlation in C60. *J. Chem. Phys.* **2011**, *135*, 11B608.

(155) Jiménez-Hoyos, C. A.; Rodrıguez-Guzmán, R.; Scuseria, G. E. Polyradical character and spin frustration in fullerene molecules: An ab initio non-collinear Hartree–Fock study. *J. Phys. Chem. A* **2014**, *118*, 9925–9940.

(156) Lee, J.; Head-Gordon, M. Distinguishing artificial and essential symmetry breaking in a single determinant: Approach and application to the C 60, C 36, and C 20 fullerenes. *Phys. Chem. Chem. Phys.* **2019**, *21*, 4763–4778.

(157) Ivanov, M. Y.; Spanner, M.; Smirnova, O. Anatomy of strong field ionization. *J. Mod. Opt.* **2005**, *52*, 165–184.

(158) Reiss, H. Limits on tunneling theories of strong-field ionization. *Phys. Rev. Lett.* **2008**, *101*, 043002.

(159) Tong, X.; Lin, C. Empirical formula for static field ionization rates of atoms and molecules by lasers in the barrier-suppression regime. *J. Phys. B* **2005**, *38*, 2593.

(160) Holmegaard, L. et al. Photoelectron angular distributions from strong-field ionization of oriented molecules. *Nat. Phys.* **2010**, *6*, 428–432.

(161) Popruzhenko, S. Keldysh theory of strong field ionization: history, applications, difficulties and perspectives. *J. Phys. B* **2014**, *47*, 204001.

(162) Shvetsov-Shilovski, N. et al. Semiclassical two-step model for strong-field ionization. *Phys. Rev. A* **2016**, *94*, 013415.

(163) Hartung, A. et al. Magnetic fields alter strong-field ionization. *Nat. Phys.* **2019**, *15*, 1222–1226.

(164) Ludwig, A.; Maurer, J.; Mayer, B.; Phillips, C.; Gallmann, L.; Keller, U. Breakdown of the dipole approximation in strong-field ionization. *Phys. Rev. Lett.* **2014**, *113*, 243001.

(165) Litvinyuk, I. V.; Lee, K. F.; Dooley, P. W.; Rayner, D. M.; Villeneuve, D. M.; Corkum, P. B. Alignment-dependent strong field ionization of molecules. *Phys. Rev. Lett.* **2003**, *90*, 233003.

(166) Mairesse, Y. et al. High harmonic spectroscopy of multichannel dynamics in strong-field ionization. *Phys. Rev. Lett.* **2010**, *104*, 213601.

(167) Schafer, K. J.; Kulander, K. C. High harmonic generation from ultrafast pump lasers. *Phys. Rev. Lett.* **1997**, *78*, 638.

(168) Christov, I. P.; Murnane, M. M.; Kapteyn, H. C. High-harmonic generation of attosecond pulses in the "single-cycle" regime. *Phys. Rev. Lett.* **1997**, *78*, 1251.

(169) Itatani, J.; Zeidler, D.; Levesque, J.; Spanner, M.; Villeneuve, D.; Corkum, P. Controlling high harmonic generation with molecular wave packets. *Phys. Rev. Lett.* **2005**, *94*, 123902.

(170) Ghimire, S.; Reis, D. A. High-harmonic generation from solids. *Nat. Phys.* **2019**, *15*, 10–16.

(171) Krausz, F.; Ivanov, M. Attosecond physics. *Rev. Mod. Phys.* **2009**, *81*, 163.

(172) Ciappina, M. F. et al. Attosecond physics at the nanoscale. *Rep. Prog. Phys.* **2017**, *80*, 054401.

(173) Scrinzi, A.; Ivanov, M. Y.; Kienberger, R.; Villeneuve, D. M. Attosecond physics. *J. Phys. B* **2005**, *39*, R1.

(174) Drescher, M.; Krausz, F. Attosecond physics: facing the wave–particle duality. *J. Phys. B* **2005**, *38*, S727.

(175) Stopkowicz, S.; Gauss, J.; Lange, K. K.; Tellgren, E. I.; Helgaker, T. Coupled-cluster theory for atoms and molecules in strong magnetic fields. *J. Chem. Phys.* **2015**, *143*, 074110.

(176) Pathak, H.; Sato, T.; Ishikawa, K. L. Time-dependent optimized coupled-cluster method for multielectron dynamics. III. A second-order many-body perturbation approximation. *J. Chem. Phys.* **2020**, *153*, 034110.

(177) Jagau, T.-C. Coupled-cluster treatment of molecular strong-field ionization. *J. Chem. Phys.* **2018**, *148*, 204102.

(178) Jagau, T.-C. Investigating tunnel and above-barrier ionization using complex-scaled coupled-cluster theory. *J. Chem. Phys.* **2016**, *145*, 204115.

(179) Krause, P.; Sonk, J. A.; Schlegel, H. B. Strong field ionization rates simulated with time-dependent configuration interaction and an absorbing potential. *J. Chem. Phys.* **2014**, *140*, 174113.

(180) Krause, P.; Schlegel, H. B. Strong-field ionization rates of linear polyenes simulated with time-dependent configuration interaction with an absorbing potential. *J. Chem. Phys.* **2014**, *141*, 174104.

(181) Kulander, K. C. Time-dependent Hartree-Fock theory of multiphoton ionization: Helium. *Phys. Rev. A* **1987**, *36*, 2726–2738.

(182) Kvaal, S. Ab initio quantum dynamics using coupled-cluster. *J. Chem. Phys.* **2012**, *136*, 194109.

(183) Pedersen, T. B.; Kvaal, S. Symplectic integration and physical interpretation of time-dependent coupled-cluster theory. *J. Chem. Phys.* **2019**, *150*, 144106.

(184) Kristiansen, H. E.; Schøyen, Ø. S.; Kvaal, S.; Pedersen, T. B. Numerical stability of time-dependent coupled-cluster methods for many-electron dynamics in intense laser pulses. *J. Chem. Phys.* **2020**, *152*, 071102.

(185) Ben-Asher, A.; Moiseyev, N. Complex absorbing potentials for stark resonances. *Int. J. Quantum Chem.* **2020**, *120*, e26067.

(186) Brown, J. M.; Kolesik, M. Properties of stark resonant states in exactly solvable systems. *Adv. Math. Phys.* **2015**, *2015*.

(187) Simon, B. The definition of molecular resonance curves by the method of exterior complex scaling. *Phys. Lett. A* **1979**, *71*, 211–214.

(188) Scrinzi, A.; Elander, N. A finite element implementation of exterior complex scaling for the accurate determination of resonance energies. *J. Chem. Phys.* **1993**, *98*, 3866–3875.

(189) Rescigno, T.; Baertschy, M.; Byrum, D.; McCurdy, C. Making complex scaling work for long-range potentials. *Phys. Rev. A* **1997**, *55*, 4253.

(190) He, F.; Ruiz, C.; Becker, A. Absorbing boundaries in numerical solutions of the time-dependent Schrödinger equation on a grid using exterior complex scaling. *Phys. Rev. A* **2007**, *75*, 053407.

(191) Rom, N.; Engdahl, E.; Moiseyev, N. Tunneling rates in bound systems using smooth exterior complex scaling within the framework of the finite basis set approximation. *J. Chem. Phys.* **1990**, *93*, 3413–3419.

(192) Morgan, J.; Simon, B. The calculation of molecular resonances by complex scaling. *J. Phys. B* **1981**, *14*, L167.

(193) McCurdy, C. W.; Martın, F. Implementation of exterior complex scaling in B-splines to solve atomic and molecular collision problems. *J. Phys. B* **2004**, *37*, 917–936.

(194) Telnov, D. A.; Sosnova, K. E.; Rozenbaum, E.; Chu, S.-I. Exterior complex scaling method in time-dependent density-functional theory: Multiphoton ionization and high-order-harmonic generation of Ar atoms. *Phys. Rev. A* **2013**, *87*, 053406.

(195) Sommerfeld, T.; Ehara, M. Complex Absorbing Potentials with Voronoi Isosurfaces Wrapping Perfectly around Molecules. *J. Chem. Theory Comput.* **2015**, *11*, PMID: 26574253, 4627–4633.

(196) Zuev, D. et al. Complex absorbing potentials within EOM-CC family of methods: Theory, implementation, and benchmarks. *J. Chem. Phys.* **2014**, *141*, 024102.

(197) Lefebvre, R.; Sindelka, M.; Moiseyev, N. Resonance positions and lifetimes for flexible complex absorbing potentials. *Phys. Rev. A* **2005**, *72*, 052704.

(198) Santra, R. Why complex absorbing potentials work: A discrete-variable-representation perspective. *Phys. Rev. A* **2006**, *74*, 034701.

(199) Vibok, A.; Balint-Kurti, G. Parametrization of complex absorbing potentials for time-dependent quantum dynamics. *J. Phys. Chem.* **1992**, *96*, 8712–8719.

(200) Jagau, T.-C.; Zuev, D.; Bravaya, K. B.; Epifanovsky, E.; Krylov, A. I. A fresh look at resonances and complex absorbing potentials: density matrix-based approach. *J. Phys. Chem. Lett.* **2014**, *5*, 310–315.

(201) Santra, R.; Cederbaum, L. S. Complex absorbing potentials in the framework of electron propagator theory. I. General formalism. *J. Chem. Phys.* **2002**, *117*, 5511–5521.

(202) Feuerbacher, S.; Sommerfeld, T.; Santra, R.; Cederbaum, L. S. Complex absorbing potentials in the framework of electron propagator theory. II. Application to temporary anions. *J. Chem. Phys.* **2003**, *118*, 6188–6199.

(203) Muga, J.; Palao, J.; Navarro, B.; Egusquiza, I. Complex absorbing potentials. *Phys. Rep.* **2004**, *395*, 357–426.

(204) Manthe, U. A time-dependent discrete variable representation for (multiconfiguration) Hartree methods. *J. Chem. Phys.* **1996**, *105*, 6989–6994.

(205) Sánchez, I.; Martın, F. Representation of the electronic continuum of withB-spline basis. *J. Phys. B* **1997**, *30*, 679–692.

(206) Martın, F. Ionization and dissociation using B-splines: photoionization of the hydrogen molecule. *J. Phys. B* **1999**, *32*, R197–R231.

(207) Bachau, H.; Cormier, E.; Decleva, P.; Hansen, J. E.; Martın, F. Applications ofB-splines in atomic and molecular physics. *Rep. Prog. Phys.* **2001**, *64*, 1815–1943.

(208) Littlejohn, R. G.; Cargo, M.; Carrington Jr, T.; Mitchell, K. A.; Poirier, B. A general framework for discrete variable representation basis sets. *J. Chem. Phys.* **2002**, *116*, 8691–8703.

(209) Tao, L.; McCurdy, C.; Rescigno, T. Grid-based methods for diatomic quantum scattering problems: A finite-element discrete-variable representation in prolate spheroidal coordinates. *Phys. Rev. A* **2009**, *79*, 012719.

(210) Yip, F. L.; McCurdy, C. W.; Rescigno, T. N. Hybrid Gaussian–discrete-variable representation for describing molecular double-ionization events. *Phys. Rev. A* **2020**, *101*, 063404.

(211) Luppi, E.; Head-Gordon, M. Computation of high-harmonic generation spectra of H-2 and N-2 in intense laser pulses using quantum chemistry methods and time-dependent density functional theory. *Mol. Phys.* **2012**, *110*, 909–923.

(212) Landau, L. D.; Lifshitz, E. M. In *Quantum Mechanics (Third Edition)*, LANDAU, L., LIFSHITZ, E., Eds., Third Edition; Pergamon: 1977, pp 133–163.

(213) Hernández Vera, M.; Jagau, T.-C. Resolution-of-the-identity second-order Møller–Plesset perturbation theory with complex basis functions: Benchmark calculations and applications to strong-field ionization of polyacenes. *J. Chem. Phys.* **2020**, *152*, 174103.

(214) Dunning Jr, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(215) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* **1992**, *96*, 6796–6806.

(216) Stewart, R. F. Small Gaussian Expansions of Slater-Type Orbitals. *J. Chem. Phys.* **1970**, *52*, 431–438.

(217) Hait, D.; Head-Gordon, M. How accurate are static polarizability predictions from density functional theory? An assessment over 132 species at equilibrium geometry. *Phys. Chem. Chem. Phys.* **2018**, *20*, 19800–19810.

(218) Simons, J. Resonance state lifetimes from stabilization graphs. *J. Chem. Phys.* **1981**, *75*, 2465–2467.

(219) Carlson, B. J.; Falcetta, M. F.; Slimak, S. R.; Jordan, K. D. A Fresh Look at the Role of the Coupling of a Discrete State with a Pseudocontinuum State in the Stabilization Method for Characterizing Metastable States. *J. Phys. Chem. Lett.* **2021**, *12*, PMID: 33481599, 1202–1206.

(220) Shao, Y. et al. Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* **2015**, *113*, 184–215.

(221) Fornace, M. E.; Lee, J.; Miyamoto, K.; Manby, F. R.; Miller III, T. F. Embedded mean-field theory. *J. Chem. Theory Comput.* **2015**, *11*, 568–580.

(222) Ding, F.; Manby, F. R.; Miller III, T. F. Embedded mean-field theory with block-orthogonalized partitioning. *J. Chem. Theory Comput.* **2017**, *13*, 1605–1615.

(223) Small, D. W.; Sundstrom, E. J.; Head-Gordon, M. A simple way to test for collinearity in spin symmetry broken wave functions: General theory and application to generalized Hartree Fock. *J. Chem. Phys.* **2015**, *142*, 094112.

(224) Goings, J. J.; Ding, F.; Frisch, M. J.; Li, X. Stability of the complex generalized Hartree-Fock equations. *J. Chem. Phys.* **2015**, *142*, 154109.

(225) Jiménez-Hoyos, C. A.; Henderson, T. M.; Scuseria, G. E. Generalized Hartree–Fock Description of Molecular Dissociation. *J. Chem. Theory Comput.* **2011**, *7*, PMID: 26605457, 2667–2674.

(226) Stetina, T. F.; Sun, S.; Williams-Young, D. B.; Li, X. Modeling magneto-Photoabsorption using time-dependent complex generalized Hartree-Fock. *ChemPhotoChem* **2019**, *3*, 739–746.

(227) Yamaki, D.; Shigeta, Y.; Yamanaka, S.; Nagao, H.; Yamaguchi, K. Generalized spin orbital calculations of spin-frustrated molecules. *Int. J. Quantum Chem.* **2001**, *84*, 546–551.

(228) Ostlund, N. S. Complex and Unrestricted Hartree-Fock Wavefunctions. *J. Chem. Phys.* **1972**, *57*, 2994–2997.

(229) Small, D. W.; Sundstrom, E. J.; Head-Gordon, M. Restricted Hartree Fock using complex-valued orbitals: A long-known but neglected tool in electronic structure theory. *J. Chem. Phys.* **2015**, *142*, 024104.

(230) Lee, J.; Head-Gordon, M. Two single-reference approaches to singlet biradicaloid problems: Complex, restricted orbitals and approximate spin-projection combined with regularized orbital-optimized Møller-Plesset perturbation theory. *J. Chem. Phys.* **2019**, *150*, 244106.

(231) Lee, J.; Bertels, L. W.; Small, D. W.; Head-Gordon, M. Kohn-Sham Density Functional Theory with Complex, Spin-Restricted Orbitals: Accessing a New Class of Densities without the Symmetry Dilemma. *Phys. Rev. Lett.* **2019**, *123*, 113001.

(232) Min, S. K.; Cho, Y.; Kim, K. S. Efficient electron dynamics with the planewave-based real-time time-dependent density functional theory: Absorption spectra, vibronic electronic spectra, and coupled electron-nucleus dynamics. *J. Chem. Phys.* **2011**, *135*, 244112.

(233) Cunha, L. A.; Lee, J.; Hait, D.; McCurdy, C. W.; Head-Gordon, M. Exploring spin symmetry-breaking effects for static field ionization of atoms: Is there an analog to the Coulson–Fischer point in bond dissociation? *J. Chem. Phys.* **2021**, *155*, 014309.

(234) Becke, A. D. Perspective: Fifty years of density-functional theory in chemical physics. *J. Chem. Phys.* **2014**, *140*.

(235) Jones, R. O. Density functional theory: Its origins, rise to prominence, and future. *Rev. Mod. Phys.* **2015**, *87*, 897.

(236) Tozer, D. J.; Peach, M. J. Density functional theory and its applications. *Phys. Chem. Chem. Phys.* **2014**, *16*, 14333–14333.

(237) Van Mourik, T.; Bühl, M.; Gaigeot, M.-P. Density functional theory across chemistry, physics and biology, 2014.

(238) Hait, D.; Head-Gordon, M. Orbital optimized density functional theory for electronic excited states. *J. Phys. Chem. Lett.* **2021**, *12*, 4517–4529.

(239) Shirley, E. L. Ab initio inclusion of electron-hole attraction: Application to x-ray absorption and resonant inelastic x-ray scattering. *Phys. Rev. Lett.* **1998**, *80*, 794.

(240) Rohlfing, M.; Louie, S. G. Electron-hole excitations and optical spectra from first principles. *Phys. Rev. B* **2000**, *62*, 4927.

(241) Cullen, J.; Krykunov, M.; Ziegler, T. The formulation of a self-consistent constricted variational density functional theory for the description of excited states. *Chem. Phys.* **2011**, *391*, 11–18.

(242) Evangelista, F. A.; Shushkov, P.; Tully, J. C. Orthogonality constrained density functional theory for electronic excited states. *J. Phys. Chem. A* **2013**, *117*, 7378–7392.

(243) Levi, G.; Ivanov, A. V.; Jónsson, H. Variational density functional calculations of excited states via direct optimization. *J. Chem. Theory Comput.* **2020**, *16*, 6968–6982.

(244) Ramos, P.; Pavanello, M. Low-lying excited states by constrained DFT. *J. Chem. Phys* **2018**, *148*.

(245) Cheng, C.-L.; Wu, Q.; Van Voorhis, T. Rydberg energies using excited state density functional theory. *J. Chem. Phys* **2008**, *129*.

(246) Roychoudhury, S.; Sanvito, S.; O'Regan, D. D. Neutral excitation density-functional theory: an efficient and variational first-principles method for simulating neutral excitations in molecules. *Sci. Rep.* **2020**, *10*, 8947.

(247) Petersilka, M. G. U. J.; Gossmann, U. J.; Gross, E. K. U. Excitation energies from time-dependent density-functional theory. *Phys. Rev. Lett.* **1996**, *76*, 1212.

(248) Casida, M. E. In *Recent Advances In Density Functional Methods: (Part I)*; World Scientific: 1995, pp 155–192.

(249) Odelius, M. Molecular dynamics simulations of fine structure in oxygen K-edge x-ray emission spectra of liquid water and ice. *Phys. Rev. B* **2009**, *79*, 144204.

(250) Guillemin, R. et al. A review of molecular effects in gas-phase KL X-ray emission. *J. Electron Spectrosc. Relat. Phenom.* **2013**, *188*, 53–61.

(251) Roychoudhury, S. et al. Controlled experiments and optimized theory of absorption spectra of Li metal and salts. *ACS Appl. Mater. Interfaces* **2021**, *13*, 45488–45495.

(252) Pascal, T. A. et al. Finite temperature effects on the X-ray absorption spectra of lithium compounds: First-principles interpretation of X-ray Raman measurements. *J. Chem. Phys.* **2014**, *140*.

(253) De Groot, F. High-resolution X-ray emission and X-ray absorption spectroscopy. *Chem. Rev.* **2001**, *101*, 1779–1808.

(254) Bergmann, U.; Glatzel, P. X-ray emission spectroscopy. *Photosynth. Res.* **2009**, *102*, 255–266.

(255) Fouda, A. A.; Besley, N. A. Improving the predictive quality of time-dependent density functional theory calculations of the X-ray emission spectroscopy of organic molecules. *J. Comput. Chem.* **2020**, *41*, 1081–1090.

(256) Hanson-Heine, M. W.; George, M. W.; Besley, N. A. Kohn-Sham density functional theory calculations of non-resonant and resonant x-ray emission spectroscopy. *J. Chem. Phys.* **2017**, *146*.

(257) Gilbert, A. T.; Besley, N. A.; Gill, P. M. Self-consistent field calculations of excited states using the maximum overlap method (MOM). *J. Phys. Chem. A* **2008**, *112*, 13164–13171.

(258) Hirata, S.; Head-Gordon, M. Time-dependent density functional theory within the Tamm–Dancoff approximation. *Chem. Phys. Lett.* **1999**, *314*, 291–299.

(259) Ge, X.; Calzolari, A.; Baroni, S. Optical properties of anthocyanins in the gas phase. *Chem. Phys. Lett.* **2015**, *618*, 24–29.

(260) Rebolini, E.; Toulouse, J. Range-separated time-dependent density-functional theory with a frequency-dependent second-order Bethe-Salpeter correlation kernel. *J. Chem. Phys.* **2016**, *144*.

(261) Lindle, D. W.; Cowan, P.; Jach, T.; LaVilla, R.; Deslattes, R.; Perera, R. C. Polarized x-ray emission studies of methyl chloride and the chlorofluoromethanes. *Phys. Rev. A* **1991**, *43*, 2353.

(262) Yumatov, V.; Okotrub, A.; Furin, G.; Salakhutdinov, N. Electronic structure of monosubstituted benzenes and X-ray emission spectroscopy: 4. Phenol. *Russ. Chem. Bull.* **1997**, *46*, 2074–2081.

(263) Epifanovsky, E. et al. Software for the frontiers of quantum chemistry: An overview of developments in the Q-Chem 5 package. *J. Chem. Phys.* **2021**, *155*.

(264) Woon, D. E.; Dunning Jr, T. H. Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon. *The Journal of chemical physics* **1993**, *98*, 1358–1371.

(265) Roychoudhury, S.; Cunha, L. A.; Head-Gordon, M.; Prendergast, D. Changes in polarization dictate necessary approximations for modeling electronic deexcitation intensity: Application to x-ray emission. *Phys. Rev. B* **2022**, *106*, 075133.

(266) Liang, Y.; Vinson, J.; Pemmaraju, S.; Drisdell, W. S.; Shirley, E. L.; Prendergast, D. Accurate x-ray spectral predictions: An advanced self-consistent-field approach inspired by many-body perturbation theory. *Phys. Rev. Lett.* **2017**, *118*, 096402.

(267) Liang, Y.; Prendergast, D. Quantum many-body effects in x-ray spectra efficiently computed using a basic graph algorithm. *Phys. Rev. B* **2018**, *97*, 205127.

(268) Liang, Y.; Prendergast, D. Taming convergence in the determinant approach for x-ray excitation spectra. *Phys. Rev. B* **2019**, *100*, 075121.

(269) Giannozzi, P. et al. QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys.: Condens. Matter* **2009**, *21*, 395502.

(270) Hybertsen, M. S.; Louie, S. G. Electron correlation in semiconductors and insulators: Band gaps and quasiparticle energies. *Phys. Rev. B* **1986**, *34*, 5390.

(271) Hedin, L.; Lundqvist, S. In *Solid state physics*; Elsevier: 1970; Vol. 23, pp 1–181.

(272) Lange, K. M.; Aziz, E. F. Electronic structure of ions and molecules in solution: a view from modern soft X-ray spectroscopies. *Chem. Soc. Rev.* **2013**, *42*, 6840–6859.

(273) Mayer, R.; Lindle, D. W.; Southworth, S.; Cowan, P. Direct determination of molecular orbital symmetry of H 2 S using polarized x-ray emission. *Phys. Rev. A* **1991**, *43*, 235.

(274) Nordgren, J.; Agren, H.; Werme, L.; Nordling, C.; Siegbahn, K. X-ray emission spectra of NH3 and N2O. *J. Phys. B* **1976**, *9*, 295.

(275) Rubensson, J.-E.; Wassdahl, N.; Brammer, R.; Nordgren, J. Local electronic structure in simple alcohols studied in X-ray emission. *J. Electron Spectrosc. Relat. Phenom.* **1988**, *47*, 131–145.

(276) Vu, T. V. et al. First-principles DFT computation and X-ray spectroscopy study of the electronic band structure and optical constants of Cu2HgGeS4. *Solid State Sci.* **2020**, *104*, 106287.

(277) Mortensen, D. et al. Benchmark results and theoretical treatments for valence-to-core x-ray emission spectroscopy in transition metal compounds. *Phys. Rev. B* **2017**, *96*, 125136.

(278) Hong, W. T.; Stoerzinger, K. A.; Moritz, B.; Devereaux, T. P.; Yang, W.; Shao-Horn, Y. Probing LaMO3 metal and oxygen partial density of states using X-ray emission, absorption, and photoelectron spectroscopy. *J. Phys. Chem. C* **2015**, *119*, 2063–2072.

(279) Grad, G. B.; González, E. R.; Torres-Dıaz, J.; Bonzi, E. V. A DFT study of ZnO, Al2O3 and SiO2; combining X-ray spectroscopy, chemical bonding and Wannier functions. *J. Phys. Chem. Solids* **2022**, *168*, 110788.

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

(281) Dreuw, A.; Fleming, G. R.; Head-Gordon, M. Charge-transfer state as a possible signature of a zeaxanthin- chlorophyll dimer in the non-photochemical quenching process in green plants. *J. Phys. Chem. B* **2003**, *107*, 6500–6503.

(282) Dreuw, A.; Head-Gordon, M. 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.* **2004**, *126*, 4007–4016.

(283) Yuhas, B. D.; Fakra, S.; Marcus, M. A.; Yang, P. Probing the local coordination environment for transition metal dopants in zinc oxide nanowires. *Nano Lett.* **2007**, *7*, 905–909.

(284) Pollock, C. J.; DeBeer, S. Insights into the geometric and electronic structure of transition metal centers from valence-to-core X-ray emission spectroscopy. *Acc. Chem. Res.* **2015**, *48*, 2967–2975.

(285) Westre, T. E.; Kennepohl, P.; DeWitt, J. G.; Hedman, B.; Hodgson, K. O.; Solomon, E. I. A multiplet analysis of Fe K-edge 1s→ 3d pre-edge features of iron complexes. *J. Am. Chem. Soc.* **1997**, *119*, 6297–6314.

(286) Solomon, E. I.; Hedman, B.; Hodgson, K. O.; Dey, A.; Szilagyi, R. K. Ligand K-edge X-ray absorption spectroscopy: covalency of ligand–metal bonds. *Coord. Chem. Rev.* **2005**, *249*, 97–129.

(287) Kubin, M. et al. Probing the oxidation state of transition metal complexes: a case study on how charge and spin densities determine Mn L-edge X-ray absorption energies. *Chem. Sci.* **2018**, *9*, 6813–6829.

(288) Chergui, M.; Collet, E. Photoinduced structural dynamics of molecular systems mapped by time-resolved X-ray methods. *Chem. Rev.* **2017**, *117*, 11025–11065.

(289) Bhattacherjee, A.; Leone, S. R. Ultrafast X-ray Transient Absorption Spectroscopy of Gas-Phase Photochemical Reactions: A New Universal Probe of Photoinduced Molecular Dynamics. *Acc. Chem. Res.* **2018**, *51*, 3203–3211.

(290) Kraus, P. M.; Zürch, M.; Cushing, S. K.; Neumark, D. M.; Leone, S. R. The ultrafast X-ray spectroscopic revolution in chemical dynamics. *Nat. Rev. Chem.* **2018**, *2*, 82–94.

(291) Ochmann, M. et al. Light-induced radical formation and isomerization of an aromatic thiol in solution followed by time-resolved x-ray absorption spectroscopy at the sulfur K-edge. *J. Am. Chem. Soc.* **2017**, *139*, 4797–4804.

(292) Bhattacherjee, A.; Pemmaraju, C. D.; Schnorr, K.; Attar, A. R.; Leone, S. R. Ultrafast intersystem crossing in acetylacetone via femtosecond x-ray transient absorption at the carbon K-edge. *J. Am. Chem. Soc.* **2017**, *139*, 16576–16583.

(293) Runge, E.; Gross, E. K. U. Density-functional theory for time-dependent systems. *Phys. Rev. Lett.* **1984**, *52*, 997.

(294) Wenzel, J.; Wormit, M.; Dreuw, A. 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.* **2014**, *35*, 1900–1915.

(295) Lopata, K.; Van Kuiken, B. E.; Khalil, M.; Govind, N. Linear-response and real-time time-dependent density functional theory studies of core-level near-edge x-ray absorption. *J. Chem. Theo. Comput.* **2012**, *8*, 3284–3292.

(296) Besley, N. A. Modeling of the spectroscopy of core electrons with density functional theory. *WIREs Comput. Mol. Sci.* **2021**, e1527.

(297) Zhang, Y.; Biggs, J. D.; Healion, D.; Govind, N.; Mukamel, S. Core and valence excitations in resonant X-ray spectroscopy using restricted excitation window time-dependent density functional theory. *J. Chem. Phys.* **2012**, *137*, 194306.

(298) Besley, N. A. Density functional theory based methods for the calculation of X-ray spectroscopy. *Acc. Chem. Res.* **2020**, *53*, 1306–1315.

(299) Besley, N. A.; Peach, M. J.; Tozer, D. J. Time-dependent density functional theory calculations of near-edge X-ray absorption fine structure with short-range corrected functionals. *Phys. Chem. Chem. Phys.* **2009**, *11*, 10350–10358.

(300) Besley, N. A.; Asmuruf, F. A. Time-dependent density functional theory calculations of the spectroscopy of core electrons. *Physical Chemistry Chemical Physics* **2010**, *12*, 12024–12039.

(301) Blake, A. V.; Wei, H.; Donahue, C. M.; Lee, K.; Keith, J. M.; Daly, S. R. Solid energy calibration standards for P K-edge XANES: electronic structure analysis of PPh_4Br. *J. Synchrotron Radiat.* **2018**, *25*, 529–536.

(302) Martin-Diaconescu, V.; Kennepohl, P. Sulfur K-edge XAS as a probe of sulfur-centered radical intermediates. *J. Am. Chem. Soc.* **2007**, *129*, 3034–3035.

(303) Minasian, S. G. et al. Determining relative f and d orbital contributions to M–Cl covalency in MCl_6^{2-} (M= Ti, Zr, Hf, U) and $UOCl^{5-}$ using Cl K-edge X-ray absorption spectroscopy and time-dependent density functional theory. *J. Am. Chem. Soc.* **2012**, *134*, 5586–5597.

(304) DeBeer George, S.; Petrenko, T.; Neese, F. Prediction of iron K-edge absorption spectra using time-dependent density functional theory. *J. Phys. Chem. A* **2008**, *112*, 12936–12943.

(305) Attar, A. R. et al. Femtosecond x-ray spectroscopy of an electrocyclic ring-opening reaction. *Science* **2017**, *356*, 54–59.

(306) Stanton, J. F.; Bartlett, R. J. The equation of motion coupled-cluster method. A systematic biorthogonal approach to molecular excitation energies, transition probabilities, and excited state properties. *J. Chem. Phys.* **1993**, *98*, 7029–7039.

(307) Coriani, S.; Koch, H. Communication: X-ray absorption spectra and core-ionization potentials within a core-valence separated coupled cluster framework. *J. Chem. Phys.* **2015**, *143*, 181103.

(308) Peng, B.; Lestrange, P. J.; Goings, J. J.; Caricato, M.; Li, X. Energy-specific equation-of-motion coupled-cluster methods for high-energy excited states: Application to K-edge X-ray absorption spectroscopy. *J. Chem. Theo. Comput.* **2015**, *11*, 4146–4153.

(309) Frati, F. et al. Coupled cluster study of the x-ray absorption spectra of formaldehyde derivatives at the oxygen, carbon, and fluorine K-edges. *J. Chem. Phys.* **2019**, *151*, 064107.

(310) Vidal, M. L.; Feng, X.; Epifanovsky, E.; Krylov, A. I.; Coriani, S. New and efficient equation-of-motion coupled-cluster framework for core-excited and core-ionized states. *J. Chem. Theory Comput.* **2019**, *15*, 3117–3133.

(311) Carbone, J. P.; Cheng, L.; Myhre, R. H.; Matthews, D.; Koch, H.; Coriani, S. An analysis of the performance of coupled cluster methods for K-edge core excitations and ionizations using standard basis sets. *Adv. Quantum Chem.* **2019**, *79*, 241–261.

(312) Besley, N. A.; Gilbert, A. T.; Gill, P. M. W. Self-consistent-field calculations of core excited states. *J. Chem. Phys.* **2009**, *130*, 124308.

(313) Derricotte, W. D.; Evangelista, F. A. Simulation of X-ray absorption spectra with orthogonality constrained density functional theory. *Phys. Chem. Chem. Phys.* **2015**, *17*, 14360–14374.

(314) Zheng, X.; Liu, J.; Doumy, G.; Young, L.; Cheng, L. Hetero-site Double Core Ionization Energies with Sub-electronvolt Accuracy from Delta-Coupled-Cluster Calculations. *J. Phys. Chem. A* **2020**, *124*, 4413–4426.

(315) Barca, G. M.; Gilbert, A. T.; Gill, P. M. W. Simple Models for Difficult Electronic Excitations. *J. Chem. Theory Comput.* **2018**, *14*, 1501–1509.

(316) Shea, J. A.; Gwin, E.; Neuscamman, E. A generalized variational principle with applications to excited state mean field theory. *J. Chem. Theory Comput.* **2020**, *16*, 1526–1540.

(317) Ye, H.-Z.; Welborn, M.; Ricke, N. D.; Van Voorhis, T. σ-SCF: A direct energy-targeting method to mean-field excited states. *J. Chem. Phys.* **2017**, *147*, 214104.

(318) Hait, D.; Head-Gordon, M. Excited state orbital optimization via minimizing the square of the gradient: General approach and application to singly and doubly excited states via density functional theory. *J. Chem. Theory Comput.* **2020**, *16*, 1699–1710.

(319) Carter-Fenk, K.; Herbert, J. M. State-Targeted Energy Projection: A Simple and Robust Approach to Orbital Relaxation of Non-Aufbau Self-Consistent Field Solutions. *J. Chem. Theory Comput.* **2020**, *16*, 5067–5082.

(320) Grofe, A. et al. Generalization of Block-Localized Wave Function for Constrained Optimization of Excited Determinants. *J. Chem. Theory Comput.* **2020**, *17*, 277–289.

(321) Hait, D.; Head-Gordon, M. Highly Accurate Prediction of Core Spectra of Molecules at Density Functional Theory Cost: Attaining Sub-electronvolt Error from a Restricted Open-Shell Kohn–Sham Approach. *J. Phys. Chem. Lett.* **2020**, *11*, 775–786.

(322) Hait, D.; Haugen, E. A.; Yang, Z.; Oosterbaan, K. J.; Leone, S. R.; Head-Gordon, M. Accurate prediction of core-level spectra of radicals at density functional theory cost via square gradient minimization and recoupling of mixed configurations. *J. Chem. Phys.* **2020**, *153*, 134108.

(323) Garner, S. M.; Neuscamman, E. Core excitations with excited state mean field and perturbation theory. *J. Chem. Phys.* **2020**, *153*, 154102.

(324) Zhao, R. et al. Dynamic-then-static approach for core excitations of open-shell molecules. *J. Phys. Chem. Lett.* **2021**, *12*, 7409–7417.

(325) Kahk, J. M.; Michelitsch, G. S.; Maurer, R. J.; Reuter, K.; Lischner, J. Core Electron Binding Energies in Solids from Periodic All-Electron Δ-Self-Consistent-Field Calculations. *J. Phys. Chem. Lett.* **2021**, *12*, PMID: 34549969, 9353–9359.

(326) Kahk, J. M.; Lischner, J. Accurate absolute core-electron binding energies of molecules, solids, and surfaces from first-principles calculations. *Phys. Rev. Materials* **2019**, *3*, 100801.

(327) Takahashi, O. Relativistic corrections for single- and double-core excitation at the K-and L-edges from Li to Kr. *Comput. Theor. Chem.* **2017**, *1102*, 80–86.

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

(329) Bussy, A.; Hutter, J. Efficient and low-scaling linear-response time-dependent density functional theory implementation for core-level spectroscopy of large and periodic systems. *Phys. Chem. Chem. Phys.* **2021**, *23*, 4736–4746.

(330) Repisky, M. et al. Excitation energies from real-time propagation of the four-component Dirac–Kohn–Sham equation. *J. Chem. Theory Comput.* **2015**, *11*, 980–991.

(331) Halbert, L.; Vidal, M. L.; Shee, A.; Coriani, S.; Severo Pereira Gomes, A. Relativistic EOM-CCSD for Core-Excited and Core-Ionized State Energies Based on the Four-Component Dirac–Coulomb (- Gaunt) Hamiltonian. *J. Chem. Theory Comput.* **2021**, *17*, 3583–3598.

(332) Kutzelnigg, W.; Liu, W. Quasirelativistic theory equivalent to fully relativistic theory. *J. Chem. Phys.* **2005**, *123*, 241102.

(333) Jensen, H. J. A. In *The Conference Talk, REHE*, 2005.

(334) Ilias, M.; Saue, T. An Infinite-Order Relativistic Hamiltonian by a Simple One-Step Transformation. *J. Chem. Phys.* **2007**, *126*, 064102.

(335) Liu, W.; Peng, D. Exact Two-component Hamiltonians Revisited. *J. Chem. Phys.* **2009**, *131*, 031104.

(336) Li, Z.; Xiao, Y.; Liu, W. On the spin separation of algebraic two-component relativistic Hamiltonians. *J. Chem. Phys.* **2012**, *137*, 154114.

(337) Cheng, L.; Gauss, J. Analytic energy gradients for the spin-free exact two-component theory using an exact block diagonalization for the one-electron Dirac Hamiltonian. *J. Chem. Phys.* **2011**, *135*, 084114.

(338) Verma, P.; Derricotte, W. D.; Evangelista, F. A. Predicting near edge X-ray absorption spectra with the spin-free exact-two-component Hamiltonian and orthogonality constrained density functional theory. *J. Chem. Theory Comput.* **2016**, *12*, 144–156.

(339) Bagus, P. S. Self-consistent-field wave functions for hole states of some Ne-like and Ar-like ions. *Phys. Rev.* **1965**, *139*, A619.

(340) Ziegler, T.; Rauk, A.; Baerends, E. J. On the calculation of multiplet energies by the Hartree-Fock-Slater method. *Theor. Chim. Acta.* **1977**, *43*, 261–271.

(341) Niskanen, J.; Norman, P.; Aksela, H.; Ågren, H. Relativistic contributions to single and double core electron ionization energies of noble gases. *J. Chem. Phys.* **2011**, *135*, 054310.

(342) Zheng, X.; Cheng, L. Performance of Delta-Coupled-Cluster Methods for Calculations of Core-Ionization Energies of First-Row Elements. *J. Chem. Theory Comput.* **2019**, *15*, 4945–4955.

(343) Frank, I.; Hutter, J.; Marx, D.; Parrinello, M. Molecular dynamics in low-spin excited states. *J. Chem. Phys.* **1998**, *108*, 4060–4069.

(344) Kowalczyk, T.; Tsuchimochi, T.; Chen, P.-T.; Top, L.; Van Voorhis, T. Excitation energies and Stokes shifts from a restricted open-shell Kohn-Sham approach. *J. Chem. Phys.* **2013**, *138*, 164101.

(345) Hait, D.; Head-Gordon, M. Delocalization errors in density functional theory are essentially quadratic in fractional occupation number. *J. Phys. Chem. Lett.* **2018**, *9*, 6280–6288.

(346) Johnson III, R. D. NIST Computational Chemistry Comparison and Benchmark Database, NIST Standard Reference Database Number 101, Release 18. October 2016. *http://cccbdb.nist.gov/* **2015**.

(347) Groom, C. R.; Bruno, I. J.; Lightfoot, M. P.; Ward, S. C. The Cambridge structural database. *Acta Crystallogr. B: Struct. Sci., Cryst. Eng. Mater.* **2016**, *72*, 171–179.

(348) McGinnety, J. A. Cesium tetrachlorocuprate. Structure, crystal forces, and charge distribution. *J. Am. Chem. Soc.* **1972**, *94*, 8406–8413.

(349) Franck, J.; Dymond, E. Elementary processes of photochemical reactions. *Trans. Far. Soc.* **1926**, *21*, 536–542.

(350) Condon, E. A theory of intensity distribution in band systems. *Phys. Rev.* **1926**, *28*, 1182.

(351) Ågren, H.; Nordgren, J.; Selander, L.; Nordling, C.; Siegbahn, K. Multiplet structure in the high-resolution x-ray emission spectrum of neon. *J. Electron Spectrosc. Relat. Phenom.* **1978**, *14*, 27–39.

(352) Banna, M.; Wallbank, B.; Frost, D.; McDowell, C.; Perera, J. Free atom core binding energies from X-ray photoelectron spectroscopy. II. Na, K, Rb, Cs, and Mg. *J. Chem. Phys.* **1978**, *68*, 5459–5466.

(353) Bodeur, S.; Millié, P.; Nenner, I. Single-and multiple-electron effects in the Si 1s photoabsorption spectra of SiX_4 (X= H, D, F, Cl, Br, CH_3, C_2H_5, OCH_3, OC_2H_5) molecules: Experiment and theory. *Phys. Rev. A* **1990**, *41*, 252.

(354) Sodhi, R. N.; Cavell, R. G. KLL Auger and core-level (1s and 2p) photoelectron shifts in a series of gaseous phosphorus compounds. *J. Electron Spectrosc. Relat. Phenom.* **1983**, *32*, 283–312.

(355) Keski-Rahkonen, O.; Krause, M. Energies and chemical shifts of the sulphur 1s level and the KL2L3 ($^1D_2$) Auger line in H_2S, SO_2 and SF_6. *J. Electron Spectrosc. Relat. Phenom.* **1976**, *9*, 371–380.

(356) Perera, R. C.; LaVilla, R. E. Molecular x-ray spectra: S-K β emission and K absorption spectra of SCO and CS_2. *J. Chem. Phys.* **1984**, *81*, 3375–3382.

(357) Sodhi, R. N.; Cavell, R. G. KLL auger and core level (1s and 2p) photoelectron shifts in a series of gaseous sulfur compounds. *J. Electron Spectrosc. Relat. Phenom.* **1986**, *41*, 1–24.

(358) Bodeur, S.; Maréchal, J.; Reynaud, C.; Bazin, D.; Nenner, I. Chlorine K shell photoabsorption spectra of gas phase HCl and Cl_2 molecules. *Zeitschrift für Physik D Atoms, Molecules and Clusters* **1990**, *17*, 291–298.

(359) Reynaud, C. et al. Electronic properties of the SF_5Cl molecule: a comparison with SF_6. I. Photoabsorption spectra near the sulphur K and chlorine K edges. *Chem. Phys.* **1992**, *166*, 411–424.

(360) Breinig, M.; Chen, M. H.; Ice, G. E.; Parente, F.; Crasemann, B.; Brown, G. S. Atomic inner-shell level energies determined by absorption spectrometry with synchrotron radiation. *Phys. Rev. A* **1980**, *22*, 520.

(361) Ambroise, M. A.; Jensen, F. Probing Basis Set Requirements for Calculating Core Ionization and Core Excitation Spectroscopy by the Δ Self-Consistent-Field Approach. *J. Chem. Theory Comput.* **2018**, *15*, 325–337.

(362) Jensen, F. Unifying general and segmented contracted basis sets. Segmented polarization consistent basis sets. *J. Chem. Theory Comput.* **2014**, *10*, 1074–1085.

(363) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *J. Phys. Chem.* **1994**, *98*, 11623–11627.

(364) Chan, B. Assessment and development of DFT with the expanded CUAGAU-2 set of group-11 cluster systems. *Int. J. Quantum Chem.* **2021**, *121*, e26453.

(365) Hui, K.; Chai, J.-D. SCAN-based hybrid and double-hybrid density functionals from models without fitted parameters. *J. Chem. Phys.* **2016**, *144*, 044114.

(366) Becke, A. D. A new mixing of Hartree–Fock and local density-functional theories. *J. Chem. Phys.* **1993**, *98*, 1372–1377.

(367) Cavell, R. G.; Jürgensen, A. Chemical shifts in P-1s photoabsorption spectra of gaseous phosphorus compounds. *J. Electron Spectrosc. Relat. Phenom.* **1999**, *101*, 125–129.

(368) Reynaud, C. et al. Double-core ionization and excitation above the sulphur K-edge in H_2S, SO_2 and SF_6. *J. Phys. B: Atomic, Molecular and Optical Physics* **1996**, *29*, 5403.

(369) Bodeur, S.; Hitchcock, A. Inner-and valence-shell excitation of SF_4 studied by photoabsorption and electron energy loss spectroscopy. *Chem. Phys.* **1987**, *111*, 467–479.

(370) Woon, D. E.; Dunning Jr, T. H. Gaussian basis sets for use in correlated molecular calculations. V. Core-valence basis sets for boron through neon. *J. Chem. Phys.* **1995**, *103*, 4572–4585.

(371) Mennucci, B. Polarizable continuum model. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 386–404.

(372) Cances, E.; Mennucci, B.; Tomasi, J. A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics. *J. Chem. Phys.* **1997**, *107*, 3032–3041.

(373) Kunze, L.; Hansen, A.; Grimme, S.; Mewes, J.-M. PCM-ROKS for the Description of Charge-Transfer States in Solution: Singlet–Triplet Gaps with Chemical Accuracy from Open-Shell Kohn–Sham Reaction-Field Calculations. *J. Phys. Chem. Lett.* **2021**, *12*, 8470–8480.

(374) Sutherland, D.; Kasrai, M.; Bancroft, G.; Liu, Z.; Tan, K. Si L-and K-edge x-ray-absorption near-edge spectroscopy of gas-phase $Si(CH_3)_x(OCH_3)_{4-x}$: Models for solid-state analogs. *Phys. Rev. B* **1993**, *48*, 14989.

(375) Engemann, C. et al. Experimental and theoretical investigations of the X-ray absorption near edge spectra (XANES) of P_4O_6 and P_4O_6X (X = O, S, Se). *Chem. Phys.* **1997**, *221*, 189–198.

(376) Ibuki, T. et al. Total photoabsorption cross-sections of CF_3SF_5 in the C, F and S K-shell regions. *Chem. Phys. Lett.* **2004**, *392*, 303–308.

(377) Ochmann, M. et al. UV-photochemistry of the disulfide bond: Evolution of early photoproducts from picosecond X-ray absorption spectroscopy at the sulfur K-Edge. *J. Am. Chem. Soc.* **2018**, *140*, 6554–6561.

(378) DeBeer George, S.; Brant, P.; Solomon, E. I. Metal and ligand K-Edge XAS of organotitanium complexes: Metal 4p and 3d contributions to pre-edge intensity and their contributions to bonding. *J. Am. Chem. Soc.* **2005**, *127*, 667–674.

(379) McKeown, D. A.; Gan, H.; Pegg, I. L.; Stolte, W. C.; Demchenko, I. X-ray absorption studies of chlorine valence and local environments in borosilicate waste glasses. *J. Nuc. Mat.* **2011**, *408*, 236–245.

(380) Shadle, S. E.; Hedman, B.; Hodgson, K. O.; Solomon, E. I. Ligand K-edge X-ray absorption spectroscopic studies: metal-ligand covalency in a series of transition metal tetrachlorides. *J. Am. Chem. Soc.* **1995**, *117*, 2259–2272.

(381) Thom, A. J.; Head-Gordon, M. Hartree–Fock solutions as a quasidiabatic basis for nonorthogonal configuration interaction. *J. Chem. Phys.* **2009**, *131*, 124113.

(382) List, N. H.; Melin, T. R. L.; van Horn, M.; Saue, T. Beyond the electric-dipole approximation in simulations of x-ray absorption spectroscopy: Lessons from relativistic theory. *J. Chem. Phys.* **2020**, *152*, 184110.

(383) Rees, J. A. et al. Experimental and theoretical correlations between vanadium K-edge X-ray absorption and Kβ emission spectra. *J. Bio. Inor. Chem.* **2016**, *21*, 793–805.

(384) Farges, F. Chromium speciation in oxide-type compounds: application to minerals, gems, aqueous solutions and silicate glasses. *Phys. Chem. Minerals* **2009**, *36*, 463–481.

(385) Hall, E. R.; Pollock, C. J.; Bendix, J.; Collins, T. J.; Glatzel, P.; DeBeer, S. Valence-to-core-detected X-ray absorption spectroscopy: Targeting ligand selectivity. *J. Am. Chem. Soc.* **2014**, *136*, 10076–10084.

(386) Lancaster, K. M.; Finkelstein, K. D.; DeBeer, S. Kβ X-ray emission spectroscopy offers unique chemical bonding insights: revisiting the electronic structure of ferrocene. *Inorg. Chem.* **2011**, *50*, 6767–6774.

(387) Liu, W.; Borg, S. J.; Testemale, D.; Etschmann, B.; Hazemann, J.-L.; Brugger, J. Speciation and thermodynamic properties for cobalt chloride complexes in hydrothermal fluids at 35–440 C and 600 bar: an in-situ XAS study. *Geochim. Cosmochim. Acta .* **2011**, *75*, 1227–1248.

(388) DiMucci, I. M. et al. The myth of d^8 copper (III). *J. Ame. Chem. Soc.* **2019**, *141*, 18508–18520.

(389) Balabanov, N. B.; Peterson, K. A. Systematically convergent basis sets for transition metals. I. All-electron correlation consistent basis sets for the 3d elements Sc–Zn. *J. Chem. Phys.* **2005**, *123*, 064107.

(390) Yamamoto, T. Assignment of pre-edge peaks in K-edge x-ray absorption spectra of 3d transition metal compounds: electric dipole or quadrupole? *X-Ray Spectrometry: An International Journal* **2008**, *37*, 572–584.

(391) Southworth, S. H. et al. Observing pre-edge K-shell resonances in Kr, Xe, and XeF_2. *Phys. Rev. A* **2019**, *100*, 022507.

(392) Kozioł, K.; Aucar, G. A. QED effects on individual atomic orbital energies. *J. Chem. Phys.* **2018**, *148*, 134101.

(393) Wen, A.; Hitchcock, A. Inner shell spectroscopy of $(\eta^5\text{-}C_5H_5)_2TiCl_2$,$(\eta^5\text{-}C_5H_5)\,TiCl_3$, and $TiCl_4$. *Can. J. Chem.* **1993**, *71*, 1632–1644.

(394) Mitzner, R. et al. L-edge x-ray absorption spectroscopy of dilute systems relevant to metalloproteins using an x-ray free-electron laser. *J. Phys. Chem. Lett.* **2013**, *4*, 3641–3647.

(395) Hocking, R. K.; Wasinger, E. C.; de Groot, F. M.; Hodgson, K. O.; Hedman, B.; Solomon, E. I. Fe L-edge XAS studies of $K_4[Fe(CN)_6]$ and $K_3[Fe(CN)_6]$: a direct probe of back-bonding. *J. Am. Chem. Soc.* **2006**, *128*, 10442–10451.

(396) Wen, A.; Rühl, E.; Hitchcock, A. Inner-shell excitation of organoiron compounds by electron impact. *Organometallics* **1992**, *11*, 2559–2569.

(397) Cunha, L. A.; Hait, D.; Kang, R.; Mao, Y.; Head-Gordon, M. Relativistic Orbital-Optimized Density Functional Theory for Accurate Core-Level Spectroscopy. *The Journal of Physical Chemistry Letters* **2022**, *13*, PMID: 35412838, 3438–3449.

(398) Buttersack, T. et al. Photoelectron spectra of alkali metal–ammonia microjets: From blue electrolyte to bronze metal. *Science* **2020**, *368*, 1086–1091.

(399) Gessner, O.; Gühr, M. Monitoring Ultrafast Chemical Dynamics by Time-Domain X-ray Photo- and Auger-Electron Spectroscopy. *Acc. Chem. Res.* **2016**, *49*, 138–145.

(400) Roth, F. et al. Direct observation of charge separation in an organic light harvesting system by femtosecond time-resolved XPS. *Nat. Commun.* **2021**, *12*, 1196.

(401) Katayama, T. et al. Tracking multiple components of a nuclear wavepacket in photoexcited Cu(I)-phenanthroline complex using ultrafast X-ray spectroscopy. *Nat. Commun.* **2019**, *10*, 3606.

(402) Zimmermann, P. et al. Modern X-ray spectroscopy: XAS and XES in the laboratory. *Coord. Chem. Rev.* **2020**, *423*, 213466.

(403) Bauernschmitt, R.; Ahlrichs, R. Treatment of electronic excitations within the adiabatic approximation of time dependent density functional theory. *Chem. Phys. Lett.* **1996**, *256*, 454–464.

(404) Furche, F. On the density matrix based approach to time-dependent density functional response theory. *J. Chem. Phys.* **2001**, *114*, 5982–5992.

(405) Marques, M. A. L.; Gross, E. K. U. Time-dependent density functional theory. *Annu. Rev. Phys. Chem.* **2004**, *55*, 427–455.

(406) *Fundamentals of Time-Dependent Density Functional Theory*; Marques, M. A. L., Maitra, N. T., Nogueira, F. M. S., Gross, E. K. U., Rubio, A., Eds.; Lecture Notes in Physics, Vol. 837; Springer: New York, 2012.

(407) Burke, K.; Werschnik, J.; Gross, E. K. U. Time-dependent density functional theory: Past, present, and future. *J. Chem. Phys.* **2005**, *123*, 062206:1–9.

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

(409) Lestrange, P. J.; Nguyen, P. D.; Li, X. Calibration of energy-specific TDDFT for modeling K-edge XAS spectra of light elements. *J. Chem. Theory Comput.* **2015**, *11*, 2994–2999.

(410) Chantzis, A.; Kowalska, J. K.; Maganas, D.; DeBeer, S.; Neese, F. *Ab initio* wave function-based determination of element specific shifts for the efficient calculation of x-ray absorption spectra of main group elements and first row transition metals. *J. Chem. Theory Comput.* **2018**, *14*, 3686–3702.

(411) Bussy, A.; Hutter, J. First-principles correction scheme for linear-response time-dependent density functional theory calculations of core electronic states. *J. Chem. Phys.* **2021**, *155*, 034108:1–10.

(412) do Couto, P. C.; Hollas, D.; Slavíček, P. On the performance of optimally tuned range-separated hybrid functionals for x-ray absorption modeling. *J. Chem. Theory Comput.* **2015**, *11*, 3234–3244.

(413) Fransson, T.; Brumboiu, I. E.; Vidal, M. L.; Norman, P.; Coriani, S.; Dreuw, A. XABOOM: An x-ray absorption benchmark of organic molecules based on carbon, nitrogen, and oxygen 1s $\rightarrow \pi^*$ transitions. *J. Chem. Theory Comput.* **2021**, *17*, 1618–1637.

(414) Song, J.-W.; Watson, M. A.; Nakata, A.; Hirao, K. Core-excitation energy calculations with a long-range corrected hybrid exchange-correlation functional including a short-range Gaussian attenuation (LCgau-BOP). *J. Chem. Phys.* **2008**, *129*, 184113:1–9.

(415) Song, J.-W.; Watson, M. A.; Hirao, K. An improved long-range corrected hybrid functional with vanishing Hartree–Fock exchange at zero interelectronic distance (LC2gau-BOP). *J. Chem. Phys.* **2009**, *131*, 144108:1–9.

(416) Capano, G. et al. The role of Hartree–Fock exchange in the simulation of x-ray absorption spectra: A study of photoexcited $[Fe(bpy)_3]^{2+}$. *Chem. Phys. Lett.* **2013**, *580*, 179–184.

(417) Besley, N. A. Fast time-dependent density functional theory calculations of the x-ray absorption spectroscopy of large systems. *J. Chem. Theory Comput.* **2016**, *12*, 5018–5025.

(418) Roemelt, M.; Maganas, D.; DeBeer, S.; Neese, F. A combined DFT and restricted open-shell configuration interaction method including spin-orbit coupling: Application to transition metal L-edge x-ray absorption spectroscopy. *J. Chem. Phys.* **2013**, *138*, 204101:1–22.

(419) Maganas, D. et al. First principles calculations of the structure and V L-edge x-ray absorption spectra of V_2O_5 using local pair natural orbital coupled cluster theory and spin–orbit coupled configuration interaction approaches. *Phys. Chem. Chem. Phys.* **2013**, *15*, 7260–7276.

(420) Seidu, I.; Neville, S. P.; Kleinschmidt, M.; Heil, A.; Marian, C. M.; Schuurman, M. S. The simulation of X-ray absorption spectra from ground and excited electronic states using core-valence separated DFT/MRCI. *J. Chem. Phys.* **2019**, *151*, 144104.

(421) Slater, J. C.; Wood, J. H. Statistical exchange and the total energy of a crystal. *Int. J. Quantum Chem.* **1971**, *5*, 3–34.

(422) Slater, J. C. Statistical exchange-correlation in the self-consistent field. *Adv. Quantum Chem.* **1972**, *6*, 1–92.

(423) Stener, M.; Lisini, A.; Decleva, P. Density functional calculations of excitation energies and oscillator strengths for C1s $\rightarrow \pi^*$ and O1s $\rightarrow \pi^*$ excitations and ionization potentials in carbonyl containing molecules. *Chem. Phys.* **1995**, *191*, 141–154.

(424) Triguero, L.; Pettersson, L. G. M.; Ågren, H. Calculations of near-edge x-ray-absorption spectra of gas-phase and chemisorbed molecules by means of density-functional and transition-potential theory. *Phys. Rev. B* **1998**, *58*, 8097–8110.

(425) Triguero, L.; Pettersson, L. G. M.; Ågren, H. Calculations of x-ray emission spectra of molecules and surface adsorbates by means of density functional theory. *J. Phys. Chem. A* **1998**, *102*, 10599–10607.

(426) Cavalleri, M.; Odelius, M.; Nordlund, D.; Nilsson, A.; Pettersson, L. G. M. Half or full core hole in density functional theory x-ray absorption spectrum calculations of water? *Phys. Chem. Chem. Phys.* **2005**, *7*, 2854–2858.

(427) Leetmaa, M.; Ljungberg, M. P.; Lyubartsev, A.; Nilsson, A.; Pettersson, L. G. M. Theoretical approximations to x-ray absorption spectroscopy of liquid water and ice. *J. Electron Spectrosc.* **2010**, *177*, 135–157.

(428) Fransson, T.; Zhotobriukh, I.; Coriani, S.; Wikfeldt, K. T.; Norman, P.; Pettersson, L. G. M. Requirements of first-principles calculations of x-ray absorption spectra of liquid water. *Phys. Chem. Chem. Phys.* **2016**, *18*, 566–583.

(429) Michelitsch, G. S.; Reuter, K. Efficient simulation of near-edge x-ray absorption fine structure (NEXAFS) in density-functional theory: Comparison of core-level constraining approaches. *J. Chem. Phys.* **2019**, *150*, 074104.

(430) Jana, D.; Bandyopadhyay, B.; Mukherjee, D. Development and applications of a relaxation-inducing cluster expansion theory for treating strong relaxation and differential correlation effects. *Theor. Chem. Acc.* **1999**, *102*, 317–327.

(431) Rankine, C. D.; Penfold, T. J. Progress in the theory of x-ray spectroscopy: From quantum chemistry to machine learning and ultrafast dynamics. *J. Phys. Chem. A* **2021**, *125*, 4276–4293.

(432) Maitra, N. T. Undoing static correlation: Long-range charge transfer in time-dependent density-functional theory. *J. Chem. Phys.* **2005**, *122*, 234104:1–6.

(433) Hait, D.; Oosterbaan, K. J.; Carter-Fenk, K.; Head-Gordon, M. Computing X-ray absorption spectra from linear-response particles atop optimized holes. *J. Chem. Phys.* **2022**, *156*, 201104:1–9.

(434) Duchemin, I.; Deutsch, T.; Blase, X. Short-Range to Long-Range Charge-Transfer Excitations in the Zincbacteriochlorin-Bacteriochlorin Complex: A Bethe-Salpeter Study. *Phys. Rev. Lett.* **2012**, *109*, 167801:1–6.

(435) Jin, Y.; Yang, W. Excitation energies from single-particle Green's function with the *GW* approximation. *J. Phys. Chem. A* **2019**, *123*, 3199–3204.

(436) Fransson, T.; Dreuw, A. Simulating x-ray emission spectroscopy with algebraic diagrammatic construction schemes for the polarization propagator. *J. Chem. Theory Comput.* **2019**, *15*, 546–556.

(437) Dreuw, A.; Fransson, T. Using core-hole reference states for calculating X-ray photoelectron and emission spectra. *Phys. Chem. Chem. Phys.* **2022**, *24*, 11259–11267.

(438) Wilhelm, J.; Seewald, P.; Golze, D. Low-scaling *GW* with benchmark accuracy and application to phosphorene nanosheets. *J. Chem. Theory Comput.* **2021**, *17*, 1662–1677.

(439) Bene, J. E. D.; Ditchfield, R.; Pople, J. A. Self-Consistent Molecular Orbital Methods. X. Molecular Orbital Studies of Excited States with Minimal and Extended Basis Sets. *J. Chem. Phys.* **1971**, *55*, 2236–2241.

(440) Foresman, J. B.; Head-Gordon, M.; Pople, J. A.; Frisch, M. J. Toward a systematic molecular orbital theory for excited states. *J. Phys. Chem.* **1992**, *96*, 135–149.

(441) Ågren, H.; Carravetta, V.; Vahtras, O.; Pettersson, L. G. M. Direct, atomic orbital, static exchange calculations of photoabsorption spectra of large molecules and clusters. *Chem. Phys. Lett.* **1994**, *222*, 75–81.

(442) Ågren, H.; Carravetta, V.; Vahtras, O.; Pettersson, L. G. Direct SCF direct static-exchange calculations of electronic spectra. *Theo. Chem. Acc.* **1997**, *97*, 14–40.

(443) Oosterbaan, K. J.; White, A. F.; Head-Gordon, M. Non-orthogonal configuration interaction with single substitutions for the calculation of core-excited states. *J. Chem. Phys.* **2018**, *149*, 044116.

(444) Oosterbaan, K. J.; White, A. F.; Head-Gordon, M. Non-orthogonal configuration interaction with single substitutions for core-excited states: An extension to doublet radicals. *J. Chem. Theory Comput.*

(445) Oosterbaan, K. J.; White, A. F.; Hait, D.; Head-Gordon, M. Generalized single excitation configuration interaction: An investigation into the impact of the inclussion of non-orthogonality on the calculation of core-excited states. *Phys. Chem. Chem. Phys.* **2020**, *22*, 8182–8192.

(446) Liu, J.; Hättig, C.; Höfener, S. Analytical nuclear gradients for electron-attached and electron-detached states for the second-order algebraic diagrammatic construction scheme combined with frozen-density embedding. *J. Chem. Phys.* **2020**, *152*, 174109:1–15.

(447) Tsudeda, T.; Song, J.-W.; Suzuki, S.; Hirao, K. On Koopmans' theorem in density functional theory. *J. Chem. Phys.* **2010**, *133*, 174101:1–9.

(448) Fischer, S. A.; Cramer, C. J.; Govind, N. Excited state absorption from real-time time-dependent density functional theory. *J. Chem. Theory Comput.* **2015**, *11*, 4294–4303.

(449) Bowman, D. N.; Asher, J. C.; Fischer, S. A.; Cramer, C. J.; Govind, N. Excited-state absorption in tetrapyridy. porphyrins: Comparing real-time and quadratic-response time-dependent density functional theory. *Phys. Chem. Chem. Phys.* **2017**, *19*, 27452–27462.

(450) Cavaletto, S. M.; Nascimento, D. R.; Zhang, Y.; Govind, N.; Mukamel, S. Resonant stimulated x-ray Raman spectroscopy of mixed-valence manganese complexes. *J. Phys. Chem. Lett.* **2021**, *12*, 5925–5931.

(451) Liekhus-Schmaltz, C. E. et al. Ultrafast x-ray pump x-ray probe transient absorption spectroscopy: A computational study and proposed experiment probing core–valence electronic correlations in solvated complexes. *J. Chem. Phys.* **2021**, *154*, 214107:1–9.

(452) Loe, C. M.; Liekhus-Schmaltz, C.; Govind, N.; Khalil, M. Spectral signatures of ultrafast excited-state intramolecular proton transfer from computational multi-edge transient x-ray absorption spectroscopy. *J. Phys. Chem. Lett.* **2021**, *12*, 9840–9847.

(453) Maitra, N. T. Charge transfer in time-dependent density functional theory. *Journal of Physics: Condensed Matter* **2017**, *29*, 423001.

(454) Imamura, Y.; Nakai, H. Analysis of self-interaction correction for describing core excited states. *Int. J. Quantum Chem.* **2007**, *107*, 23–29.

(455) Goscinski, O.; Weiner, B. The Role of Algebraic Formulations of Approximate Green's Functions for Systems With a Finite Number of Electrons. *Phys. Scr.* **1980**, *21*, 385.

(456) Weiner, B.; Goscinski, O. Self-consistent approximation to the polarization propagator. *Int. J. Quantum Chem.* **1980**, *18*, 1109–1131.

(457) Prasad, M. D.; Pal, S.; Mukherjee, D. Some aspects of self-consistent propagator theories. *Phys. Rev. A* **1985**, *31*, 1287–1298.

(458) Datta, B.; Mukhopadhyay, D.; Mukherjee, D. Consistent propagator theory based on the extended coupled-cluster parametrization of the ground state. *Phys. Rev. A* **1993**, *47*, 3632–3648.

(459) Carter-Fenk, K.; Head-Gordon, M. On the choice of reference orbitals for linear-response calculations of solution-phase K-edge X-ray absorption spectra. *pccp* **2022**, *24*, 26170–26179.

(460) Bourne Worster, S.; Feighan, O.; Manby, F. R. Reliable transition properties from excited-state mean-field calculations. *J. Chem. Phys.* **2021**, *154*, 124106.

(461) Anderson, L. N.; Oviedo, M. B.; Wong, B. M. Accurate electron affinities and orbital energies of anions from a nonempirically tuned range-separated density functional theory approach. *J. Chem. Theory Comput.* **2017**, *13*, 1656–1666.

(462) Zhou, B.; Hu, Z.; Jiang, Y.; He, X.; Sun, Z.; Sun, H. Benchmark study of ionization potentials and electron affinities of armchair single-walled carbon nanotubes using density functional theory. *J. Phys.: Condens. Matter* **2018**, *30*, 215501.

(463) Schirmer, J. et al. K-shell excitation of the water, ammonia, and methane molecules using high-resolution photoabsorption spectroscopy. *Phys. Rev. A* **1993**, *47*, 1136.

(464) Grimm, S.; Nonnenberg, C.; Frank, I. Restricted open-shell Kohn–Sham theory for π–π transitions: I. Polyenes, cyanines, and protonated imines. *J. Chem. Phys.* **2003**, *119*, 11574–11584.

(465) Billeter, S. R.; Egli, D. Calculation of nonadiabatic couplings with restricted open-shell Kohn-Sham density-functional theory. *J. Chem. Phys.* **2006**, *125*, 224103:1–18.

(466) Filatov, M.; Shaik, S. Application of spin-restricted open-shell Kohn–Sham method to atomic and molecular multiplet states. *J. Chem. Phys.* **1999**, *110*, 116–125.

(467) Friedrichs, J.; Damianos, K.; Frank, I. Solving restricted open-shell equations in excited state molecular dynamics simulations. *Chem. Phys.* **2008**, *347*, 17–24.

(468) Von Barth, U. Local-density theory of multiplet structure. *Phys. Rev. A* **1979**, *20*, 1693–1703.

(469) Görling, A. Symmetry in density-functional theory. *Phys. Rev. A* **1993**, *47*, 2783–2799.

(470) Levy, M.; Nagy, Á. Variational Density-Functional Theory for an Individual Excited State. *Phys. Rev. Lett.* **1999**, *83*, 4361–4364.

(471) Sodhi, R.; Brion, C. High resolution carbon 1s and valence shell electronic excitation spectra of trans-1,3-butadiene and allene studied by electron energy loss spectroscopy. *J. Electron Spectrosc. Relat. Phenom.* **1985**, *37*, 1–21.

(472) Carter-Fenk, K.; Cunha, L. A.; Arias-Martinez, J. E.; Head-Gordon, M. Electron-Affinity Time-Dependent Density Functional Theory: Formalism and Applications to Core-Excited States. *J. Phys. Chem. Lett.* **2022**, *13*, PMID: 36215404, 9664–9672.

(473) Ye, H.-Z.; Berkelbach, T. C. Correlation-consistent Gaussian basis sets for solids made simple. *J. Chem. Theory Comput.* **2022**, *18*, 1595–1606.

(474) Cui, Z.-H. et al. Ab initio quantum many-body description of superconducting trends in the cuprates. *arXiv preprint arXiv:2306.16561* **2023**.

(475) Hammes-Schiffer, S. Exploring proton-coupled electron transfer at multiple scales. *Nat. Comp. Sci.* **2023**, *3*, 291–300.

(476) Hammes-Schiffer, S. Nuclear–electronic orbital methods: Foundations and prospects. *J. Chem. Phys.* **2021**, *155*.

(477) Ruggenthaler, M.; Flick, J.; Pellegrini, C.; Appel, H.; Tokatly, I. V.; Rubio, A. Quantum-electrodynamical density-functional theory: Bridging quantum optics and electronic-structure theory. *Phys. Rev. A* **2014**, *90*, 012508.

(478) Hoffmann, N. M.; Appel, H.; Rubio, A.; Maitra, N. T. Light-matter interactions via the exact factorization approach. *Eur. Phys. J. B* **2018**, *91*, 1–14.

(479) Haugland, T. S.; Ronca, E.; Kjønstad, E. F.; Rubio, A.; Koch, H. Coupled cluster theory for molecular polaritons: Changing ground and excited states. *Phys. Rev. X* **2020**, *10*, 041043.

(480) Ruggenthaler, M.; Tancogne-Dejean, N.; Flick, J.; Appel, H.; Rubio, A. From a quantum-electrodynamical light–matter description to novel spectroscopies. *Nat. Rev. Chem.* **2018**, *2*, 1–16.

(481) Bohm, D.; Pines, D. A collective description of electron interactions: III. Coulomb interactions in a degenerate electron gas. *Phys. Rev.* **1953**, *92*, 609.

(482) Bardeen, J.; Pines, D. Electron-phonon interaction in metals. *Phys. Rev.* **1955**, *99*, 1140.

(483) Wouters, S.; Jiménez-Hoyos, C. A.; Sun, Q.; Chan, G. K.-L. A practical guide to density matrix embedding theory in quantum chemistry. *J. Chem. Theory Comput.* **2016**, *12*, 2706–2719.

(484) Lacombe, L.; Maitra, N. T. Embedding via the exact factorization approach. *Phys. Rev. Lett.* **2020**, *124*, 206401.

(485) Sheng, N.; Vorwerk, C.; Govoni, M.; Galli, G. Green's function formulation of quantum defect embedding theory. *J. Chem. Theory Comput.* **2022**, *18*, 3512–3522.