Cooking core spectroscopy with the finest ingredients: Computational recipes based on
orbital-optimized references

by

Juan Esteban Arias Martinez

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 Jeffrey Neaton

Fall 2023

Cooking core spectroscopy with the finest ingredients: Computational recipes based on orbital-optimized references

Abstract

Cooking core spectroscopy with the finest ingredients: Computational recipes based on orbital-optimized references

by

Juan Esteban Arias Martinez

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor Martin Head-Gordon, Chair

The development of predictive computational methods to simulate, and thus interpret the phenomena relevant to core spectroscopies is a vibrant and flourishing field. Until recently, computational core spectroscopy was under-developed relative to ground-state and even valence excited-state quantum chemistry. By the appeal of its broad range of modern experimental applications, this emerging sub-field of quantum chemistry is a fertile landscape for research. The work presented here grew in this landscape, and the different findings are bound together by the common theme of using of orbital-optimized references for the description of core excited states. After laying a foundation in Chapter 1, Chapter 2 explores schemes for the use of orbital-optimized references as ingredients for correlated calculations with the aim of providing high quality predictions of the energy required to induce a core excitation. The schemes proposed rely on single-reference coupled-cluster singles and doubles - long lauded for being the most affordable variant within the accurate coupled cluster formalism - and achieves a statistical performance on the order of 0.2 eV for core excitations in small organic molecules. Chapter 3 designs a model constructed within a generalized single-excitation CI framework, relying on orbital-optimized references, to describe the exotic core excited states present in ultra-fast pump-probe core spectroscopy experiments. Specifically, these are core excited states atop valence excited states used to discern the electronic dynamics during photochemical processes. As an efficient yet accurate zeroth order model enjoying explicit spin adaptation, we demonstrate the utility of our model by simulating an ultra-fast time-resolved core absorption spectra of acetylacetone that requires small shifts in energy to align with experiment, and is free from artifacts due to spin contamination. Chapter 4 uses the aforementioned model for an exciting photochemical application on $Fe(CO)_5$, a textbook organometallic compound used to understand metal-ligand bonding in these type of complexes. With our valence-core excited state model, we help uncover for the first time the spectroscopic signatures of the metal-centered excited states of $Fe(CO)_5$ during photodissociation. Finally, Chapter 5 offers some concluding remarks and future outlook.

Te dedico esta obra a ti Juanes. Eres un berraco.

Contents

List of Figures

List of Tables

Acknowledgments

Had it not been for the handful of friends, family, mentors, and angels scattered throughout the road, I would not be writing these words. The pandemic and isolation was awful. Really awful. Thanks are due to Ellen because she was truly a rock star when it came to keeping me in the program and I am glad she was. Thanks are due to Dip as well, for making isolation a bit more fun with films and video games, and helping me through the rough times. Thank you to Leo, for his friendship and support as well during the days when Gilman 42 would have to convene virtually.

I must tip my hat off to Martin. He saw me at my worst during the pandemic and his support and understanding were essential. Martin has a superpower with his students when they're learning, or at least he had with me: he gets down to your level and makes you feel like you're smart. In part because of his superpower, my PhD has been like an adventure with quests where we uncover the interesting treasures in the land of electronic structure and core spectroscopy. No one is perfect, and Martin handled my criticism with grace when I needed to share it. All in all, we managed to pull through the pandemic and turn my PhD around into a successful enterprise. Thank you for letting a kid from Florida dream about quantum chemistry, Martin.

Speaking of Florida, let me also give a shout out to the mentors and advisors from UCF that helped me get here. I'd like to acknowledge Dr. Torres, Dr. Hernandez, and Dr. Harper for making a chemist out of me. They successfully nurtured my excitement about science and their mentor-ship and faith was indispensable to getting me into graduate school. The truth is I'd likely not be here without Michael, Natalia, Ms. Arlene and the rest of the McNair family. The deep respect and appreciation I feel for them are beyond my English at the moment. Lets get back to California now.

My gratitude to the Gilman core mafia. Kevin (C-F.) shares Martin's superpower and it has been a lot of fun to work with them and bounce ideas around. Without a fault, when trying to figure out a problem, I'd always realize that the answer lay hidden in something Leo brought up two months ago. I promise I'm getting better, Leo! Work aside, the Gilman core mafia's friendship has made my time in Berkeley more enjoyable through cat-sitting, nature, and dinners. Hamlin, the youngest member of the GCM, has taken numerous travels through the land of core spectroscopy with me. We've attempted, failed, learned, and succeeded. It's been an honor to work together and I'm looking forward to see where time will take him.

While the GCM where my main colleagues, I've had the pleasure of interacting with many other wonderful scientists throughout (and before) the program. I owe Brad many thanks for taking me under his wing during my first encounter with the MHG group, on Summer of 2018. Josh' great pedagogical skills were key to preparing me for my qualifying exams - it was wonderful to share an office with him for some time. Katherine led me by the hand during my first few steps into the world of core spectroscopy and, together with Leo, into the mighty QChem. After long travels therein, I took a fascinating excursion into the realm of molecules that move and Justin was an excellent companion for it. I had the good fortune (thank you Kevin) of meeting the team at Sandia National Lab (Jan, Krupa, Laura,

Chapter 1

Introduction

Core spectroscopy involves disturbing the innermost electrons of a molecule - those orbiting the nuclei most closely - with electromagnetic radiation. As a consequence of their strong localization near the nuclei, core electrons serve as a spectroscopic probe capable of providing information about the chemical environment in the vicinity of certain atoms within a molecule. Specifically, since the energetic cost of exciting core electrons is strongly dependent on the nuclei to which they belong, the wavelengths of light used to probe core electrons can be tuned to focus exclusively on atoms of a specific element. These characteristics make of core spectroscopy an excellent tool for characterization tasks across a wide variety of disciplines, ranging from detecting the elemental composition of astrophysical plasmas to uncovering the coordination environment and oxidation state of transition metal complexes relevant to catalysis. [1, 2] Fundamental studies of photochemical phenomena are another major client of core spectroscopies; this subject is of key relevance to two thirds of this thesis. Specifically, ultra-fast core spectroscopies are opening a window to uncover the nuclear and electronic dynamics of photochemical reactions by leveraging the advances in the generation of short X-ray pulses, now venturing into the attosecond (10^{-18} s) regime. [3–6]

Before delving into the main work, this Introduction serves to lay a solid foundation to discuss the relevant quantum chemistry. In Section 1.1, we arrive at the central problem of electronic structure - finding approximations to the electronic energies - by drawing from the general knowledge I've acquired from my courses in quantum mechanics and introductory textbooks. [7–10] Section 1.2 is devoted to describing a variety of relevant methods for finding approximations to the lowest electronic energy of a molecular system, i.e. the ground state energy. Section 1.3 focuses on the methods that aim to find approximations to the higher electronic energies, i.e. the excited state energies, with an emphasis on the advances specific to core excited states. Finally, Section 1.4 provides an overview of the following Chapters, as well as some of the research I was a part of but that will not be included in this thesis.

1.1 Basic premise of electronic structure

1.1.1 Introducing an electronic Hamiltonian

A quantum system is described by an associated state vector $|\Psi\rangle$ and evolves in time according to the Schrödinger equation, as dictated by the Hamiltonian operator $\hat{H}$ acting on it. [11] In atomic units, in use throughout this thesis, the Schrödinger equation reads

$$i\frac{\partial}{\partial t}|\Psi\rangle = \hat{H}|\Psi\rangle \tag{1.1}$$

For molecular systems, the Born-Oppenheimer approximation serves to separate the nuclear and electronic degrees of freedom on the basis that electronic motion occurs on a faster time scale due to their mass relative to nuclei. [12] The Born-Oppenheimer approximation allows us to define a Hamiltonian that depends exclusively on the electronic coordinates, along with an associated Hilbert space, where the nuclei act as fixed point charges. In the absence of relativistic effects and external fields, the standard Born-Oppenheimer electronic Hamiltonian employed in electronic structure reads

$$\hat{H}_e = \hat{T}_e + \hat{V}_{eN} + \hat{V}_{ee} + \hat{V}_{NN} \tag{1.2}$$

where $\hat{T}_e$ stands for the electronic kinetic energy, $\hat{V}_{eN}$ for the electron - nuclei potential energy, $\hat{V}_{ee}$ for the electron - electron potential energy , and $\hat{V}_{NN}$ for the nuclei - nuclei potential energy. For ease of discussion, both the abstract operators (with hats) and their representation in configuration space (no hats) will be used as convenient, along with their associated abstract vectors (bras and kets) and wave functions (no bras and kets). The four contributions to the electronic Hamiltonian are:

$$T_e = -\frac{1}{2}\sum_i \nabla_i^2 \tag{1.3}$$

$$V_{eN} = -\sum_{i,\,A} \frac{Z_A}{|\mathbf{r}_i - \mathbf{R}_A|} \tag{1.4}$$

$$V_{ee} = \frac{1}{2}\sum_{i,\,j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{1.5}$$

$$V_{NN} = \frac{1}{2}\sum_{A,\,B} \frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|} \tag{1.6}$$

The eigenvalues and eigenvectors of the time-independent analogue of the electronic Schrödinger equation

$$\hat{H}_e|\Psi_k\rangle = \mathcal{E}_k|\Psi_k\rangle \tag{1.7}$$

serve as powerful constructs for quantum chemistry. Since the eigenvalues $\mathcal{E}_k\{\mathbf{R}\}$ depend parametrically on the positions of the nuclei, they define potential energy surfaces (PES)

that serve as hyperdimensional landscapes for nuclei to traverse. When those traversals connect reactants and products, via intermediates and transition structures, we are simulating chemistry.

1.1.2 Born-Oppenheimer potential energy surfaces

For a given system, the wave function associated with the lowest eigenvalue Ψ_0 is defined as the electronic ground state and all others as electronic excited states. The potential energy surfaces $\mathcal{E}_n\{\mathbf{R}\}$ guide the motion of nuclei, as dictated by a nuclear Schrödinger equation

$$i\frac{\partial}{\partial t}\left|\Theta\right\rangle = (\hat{T}_N + \mathcal{E}_k)\left|\Theta\right\rangle \tag{1.8}$$

where $\hat{T}_N$ refers to the kinetic energy of the nuclei.

$$T_N = -\frac{1}{2}\sum_A \frac{\nabla_A^2}{m_A} \tag{1.9}$$

The electron-electron interaction in Eq. 1.5 prevents the Schrödinger equation to be written as a sum of one-electron operators and the all-electron wave function to be written as a product of one electron functions. This non-separability renders the electronic problem a many-body problem which cannot be solved analytically beyond two electrons, thus giving rise to the central goal of quantum chemistry: finding approximations to the electronic energies.

1.1.3 Spin in electronic wave functions

For a system with n electrons, the eigenvectors of the electronic Hamiltonian

$$\Psi_k(\mathbf{x}_1, \cdots, \mathbf{x}_n) \tag{1.10}$$

depend on the three-dimensional coordinates of each of the electrons $\mathbf{r}$ as well as their associated spin state σ. Collectively, the spatial and spin coordinates of a single electron are denoted with the symbol $\mathbf{x}$. Spin is postulated as an inherent property of fundamental particles in non-relativistic quantum mechanics that causes them to interact with external magnetic fields. The spin vector space for a single electron turns out to be two-dimensional and its state is traditionally expressed in the basis of the spin eigenstates of an electron under a uniform magnetic field pointing in the z Cartesian direction. Under these conditions, the spin of an electron is colloquially said to either "point up" or "point down" with respect to the field.

$$s_z\left|\uparrow\right\rangle = \frac{1}{2}\left|\uparrow\right\rangle \tag{1.11}$$

$$s_z\left|\downarrow\right\rangle = -\frac{1}{2}\left|\downarrow\right\rangle \tag{1.12}$$

Concrete functions can be defined from these eigenvectors.

$$\langle \sigma | \uparrow \rangle = \alpha(\sigma) \tag{1.13}$$

$$\langle \sigma | \downarrow \rangle = \beta(\sigma) \tag{1.14}$$

The intrinsic spin of a single electron confers many-electron systems with a collective spin state.

$$\vec{S} = \hat{S}_x + \hat{S}_y + \hat{S}_z \tag{1.15}$$

The individual Cartesian components of the n-electron spin operator are straightforwardly defined in terms of their one-electron analogues. For example, for the z-component,

$$\hat{S}_z(\sigma_1, \cdots, \sigma_N) = \sum_i \hat{s}_z(\sigma_1) \tag{1.16}$$

While spin does not appear explicitly in the non-relativistic electronic Hamiltonian, it has profound consequences on the structure of the electronic wave functions, as elaborated in Subsection 1.2.1. The Hamiltonian operator, the spin-squared operator, and one of its components can be taken as a set of commuting operators with a shared eigenbasis and, as a result, the electronic wave functions $|\Psi_k\rangle$ are associated with a well-defined spin state. The spectrum for $\hat{S}^2$ - a measure of the spin state of a system - is

$$\hat{S}^2 |\Psi_k\rangle = S_k(S_k + 1) |\Psi_k\rangle \tag{1.17}$$

where S_k takes on non-negative, half-integer values,

$$S_k = 0, \ \frac{1}{2}, \ \frac{1}{1}, \ \frac{3}{2}, \cdots \tag{1.18}$$

The spectrum arises from the fact that the spin operators satisfy the standard commutation relationships for an angular momentum vector operator in quantum mechanics.

$$[\hat{S}_i, \hat{S}_j] = i\hat{S}_k \epsilon_{ijk} \tag{1.19}$$

$$[\hat{S}^2, \hat{S}_i] = 0 \tag{1.20}$$

In chemistry, it is common to refer to the spin state of a system by an associated positive, integer multiplicity.

$$M_k = (2S_k + 1) \tag{1.21}$$

$$= 1, \ 2, \ 3, \ \cdots \tag{1.22}$$

If $|\Psi_k\rangle$ yields an $M_k = 1$, it is said to be a singlet state; if it yields an $M_k = 2$, it is said to be a doublet state; so forth and so on...
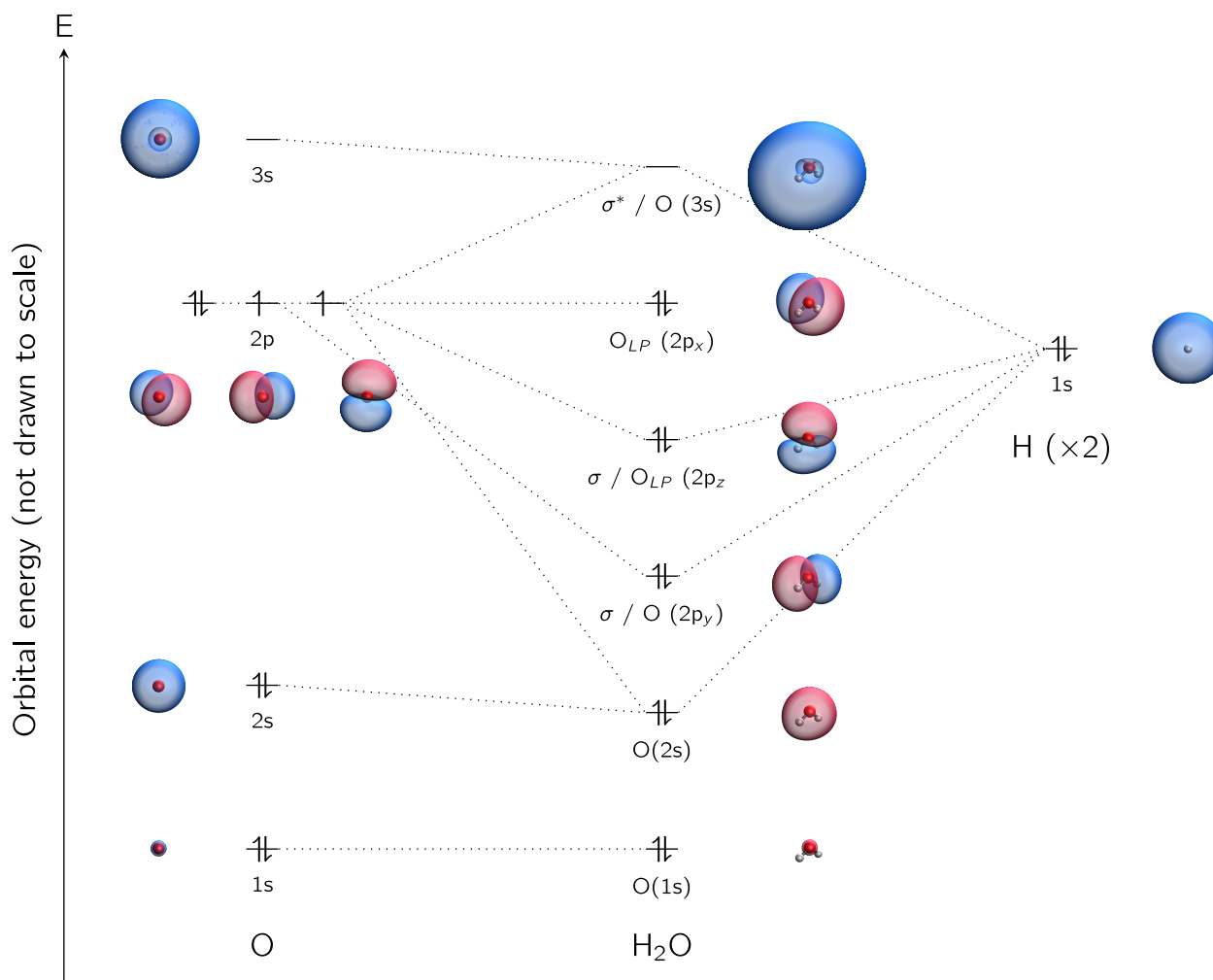
Figure 1.1: MO diagram of water. Source of orbitals: RHF / pcseg-2.

1.1.4   Molecular orbital theory

My undergraduate quantum mechanics teacher, Prof. Richard Klemm, used to say: "A picture is worth a thousand equations". Molecular orbital (MO) theory provides a conceptual framework that, while intimately tied with Hartree-Fock theory, operates in the absence of equations and mathematics and allows us to explain a broad range of phenomena in chemistry in a pictorial manner. I introduce some basic notions of qualitative MO theory to give the subsequent discussion on Hartree-Fock theory, the correlated methods, and excited state methods, more chemical shape and color.

In MO theory, electrons are conceived to inhabit MOs with origins in the interaction between the orbitals from the constituent atoms of the system, or atomic orbitals (AOs).

The degree to which two AOs will interact depends on the spatial overlap between them and their energetic proximity. AOs with more overlap that are closer in energy will mix more.

I will use water as an example to illustrate many concepts throughout this thesis, beginning with MO theory. Interacting the AOs from an oxygen atom and two hydrogen atoms at a specific distance and angle ($R_{OH} = 0.958$ Å, $\angle_{HOH} = 104.50^o$) generates a set of molecular orbitals for the water molecule. The Aufbau principle, the Pauli-exclusion principle, [13] and Hund's rule [14] provide us with a recipe to construct an MO diagram for the ground state:

- Take each of the electrons from the atoms and fill up the molecular orbitals from the bottom up in terms of energy.

- Pair up to two electrons of opposite spin within one orbital.

- When degenerate orbitals arise, fill each one with a single electron first, all of them pointing in the same direction.

The resulting diagram (Figure 1.1) illustrates how the ten electrons fill molecular orbitals of different spatial extent and orbital energies. The two electrons inhabiting the orbital with the lowest energy, largely arising from the 1s orbital of the oxygen atom (O_{1s}), are called core electrons. The remaining eight electrons inhabiting the next four orbitals belong to what is referred to as the valence space. Two electrons inhabit an orbital resembling the 2s orbital of the oxygen atom, partially fused with the two H_{1s} orbitals in what chemists know as a σ bond. The orbital with strong O_{2p_y} character also participates in a σ-bonding interaction the two H_{1s} orbitals. The last two orbitals containing electrons are more atomic in character, arising from the O_{2p_z} and O_{2p_x} orbitals, and are called the oxygen lone-pairs. Finally, MO theory states that the number of molecular orbitals must be equal to the number of the atomic orbitals that give rise to them. To compliment the bonding σ orbitals, one would expect a collection of anti-bonding σ^* orbitals but, as it turns out in water, these either mix strongly with the unoccupied 3s level of the oxygen or lie beyond it. [15] While these orbitals are not occupied, they are crucial for the quantitative theories beyond the Hartree-Fock approximation and for the description of excited states.

Most of chemistry is concerned with the electronic ground state and the interactions among valence electrons. Therefore, the field focusing on the ground state wave function and energy is vast and Section 1.2 elaborates on the relevant methods. However, the whole novelty of this thesis relates to approximate excited state wave functions and energies for molecular systems exposed to irradiation with extreme ultra-violet and soft X-ray light that promote core electrons either into empty orbitals, or remove them from the molecule altogether (core spectroscopy). Section 1.3 introduces the frameworks for excited states that relate to this work, making emphasis on their performance for core spectroscopy.

1.2 Ground state electronic structure

The electronic structure methods designed to approximate to the electronic ground state wave functions and energies, and the resulting PES generated by sampling nuclear configurations $\mathcal{E}_0(\mathbf{R})$, provide valuable chemical information. For example, the minima in $\mathcal{E}_0(\mathbf{R})$ correspond to predicted stable structures for a given set of atoms. These could be two isomers of a molecule or the product and reactant in a chemical reaction.

$$E_A = \mathcal{E}_0(\mathbf{R}_A) \tag{1.23}$$
$$E_B = \mathcal{E}_0(\mathbf{R}_B) \tag{1.24}$$

The difference between the two above represents the electronic contribution to the stability of one isomer or reactant / product over the other. Diagonalizing the nuclear Hessian evaluated at a minima

$$\left.\frac{d^2\mathcal{E}_0}{d\mathbf{Q}^2}\right|_{\mathbf{R}_A} \tag{1.25}$$

where $\mathbf{Q}$ is a vector containing the Cartesian coordinates of each nuclei weighted by its mass. For example, mass-weighed coordinates for nuclei N are obtained by carrying out the following transformation:

$$\mathbf{R}_N \longrightarrow \mathbf{Q}_N = (m_N)^{1/2}\mathbf{R}_N \tag{1.26}$$

provides the predicted vibrational frequencies of a molecule in the harmonic approximation. Half its trace provides a prediction of the vibrational zero point energy. First-order saddle points in $\mathcal{E}_0(\mathbf{R})$ connecting two minima represent the corresponding transition state for a chemical reaction. Modeling nuclear motion either classically or quantum-mechanically by means of molecular dynamics simulations provides an even more realistic picture of chemistry. Enabling these and many other tasks is the purpose of ground state electronic structure. In the following sections I introduce some of the cornerstone methods for molecular systems.

1.2.1 Hartree-Fock approximation

A standard approach in the development of electronic structure consists of proposing an ansatz (a model or an approximation) for the electronic wave function. Hartree proposed the ansatz that bears his name, Hartree product, early on. [16–18]

$$\Phi_0^{\text{H.P.}} = \chi_1(\mathbf{x}_1)\cdots\chi_N(\mathbf{x}_N) \tag{1.27}$$

$\Phi_0^{\text{H.P.}}$ is a simple product of functions χ, called spin orbitals, that depend exclusively on the coordinates of a single electron, $\mathbf{x}$. Slater and Fock soon realized that a Hartree product does not respect the anti-symmetry required by a many-fermion system, [19, 20] which results in the fact that the sign of the wave function must change on exchange of the coordinates of

two electrons. The Hartree-Fock ansatz provides a solution to the problem by proposing a wave function in the form of a Slater determinant instead. [21–23]

$$\Phi_0 = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\mathbf{x}_1) & \cdots & \chi_1(\mathbf{x}_N) \\ \vdots & \ddots & \vdots \\ \chi_N(\mathbf{x}_1) & \cdots & \chi_N(\mathbf{x}_N) \end{vmatrix} \tag{1.28}$$

Exchanging the electronic coordinates of two electrons amounts to exchanging two columns in the determinant, which results in the desired change in sign. The factor of $\frac{1}{\sqrt{N!}}$ normalizes the Hartree-Fock wave function

$$\langle \Phi_0 | \Phi_0 \rangle = 1 \tag{1.29}$$

provided all of the component spin-orbitals are orthonormal themselves

$$\langle \chi_i | \chi_j \rangle = \delta_{ij} \tag{1.30}$$

The variational principle states that the expectation value of the Hamiltonian for an approximate ground state wave function, such as the Hartree-Fock wave function, can only be too high relative to the exact value

$$\langle \Phi_0 | \hat{H} | \Phi_0 \rangle = E^{\text{H.F.}} \geq \mathcal{E}_0 \tag{1.31}$$

Subject to the constraint that the spin orbitals remain orthonormal to retain the structure of the Hartree-Fock wave function, invoking the variational principle provides a path to finding the best possible energy that the Hartree-Fock method can afford by constructing and minimizing the corresponding Langrangian

$$\delta\mathcal{L} = \delta\big(\langle \Psi^{\text{H.F.}} | \hat{H} | \Psi^{\text{H.F.}} \rangle - \sum_{pq} \lambda_{qp}(\langle \chi_p | \chi_q \rangle - \delta_{pq}) \big) = 0 \tag{1.32}$$

$$\vdots$$

$$\hat{f} | \chi_p \rangle = \varepsilon_p | \chi_p \rangle \tag{1.33}$$

The result is the venerable Hartree-Fock equation, where the spin orbitals $| \chi_p \rangle$ emerge as the eigenfunctions of the Fock operator $\hat{f}$, to be defined shortly. Hartree-Fock theory and the associated spin orbitals are, in essence, the quantitative formulation of the qualitative MO theory and its orbitals, presented in Section 1.1.4.

The Fock operator is a mean-field Hamiltonian that smears the interaction among electrons into an effective one-electron potential, $v^{\text{H.F.}}$. This potential, made up of what is referred to as the Coulomb $\hat{J}$ and exchange $\hat{K}$ contributions, together with the standard

one-electron terms $\hat{h}$ of the proper electronic Hamiltonian, constitute the Fock operator.

$$f(\mathbf{r}_1) = h(\mathbf{r}_1) + v^{\text{H.F.}}(\mathbf{r}_1) \tag{1.34}$$

$$h(\mathbf{r}_1) = -\frac{\nabla_1^2}{2} - \sum_A \frac{Z_A}{\mathbf{r}_1 - \mathbf{R}_A} \tag{1.35}$$

$$v^{\text{H.F.}}(\mathbf{r}_1) = \sum_i J_i(\mathbf{r}_1) - K_i(\mathbf{r}_1) \tag{1.36}$$

$$J_i(\mathbf{r}_1) = \left[\int d\mathbf{x}_2 \chi_i^*(\mathbf{r}_2) \mathbf{r}_{12}^{-1} \chi_i(\mathbf{r}_2) \right] \tag{1.37}$$

$$K_i(\mathbf{r}_1)\chi_p(\mathbf{x}_1) = \left[\int d\mathbf{x}_2 \chi_i^*(\mathbf{x}_2) \mathbf{r}_{12}^{-1} \chi_p(\mathbf{x}_2) \right] \chi_i(\mathbf{x}_1) \tag{1.38}$$

Note the interesting form of the exchange contribution, which is defined by its operation on an arbitrary spin orbital.

Flavors of Hartree-Fock

A practical form of the Hartree-Fock wave function requires further specifying the constraints on the orbitals themselves. In their most general form, the spin orbitals are complex-valued functions with spatial functions (also referred to as molecular orbitals) associated with independent contributions from an electron's α and β spin components. [24]

$$\chi_i(\mathbf{x}) = C_i^a \phi_i^\alpha(\mathbf{r})\alpha(\sigma) + C_i^\beta \phi_i^\beta(\mathbf{r})\beta(\sigma) \tag{1.39}$$

In fact, this form is more general than usually needed and there are only two special cases of spin orbitals relevant to this work. Both cases assume them to be real-valued and associated with either an α or a β spin function. In the unrestricted Hartree-Fock (UHF) model, all of the spin orbitals are comprised of two independent set of orbitals, one α and one β.

$$\{ \chi_1 \cdots \chi_N \} = \begin{cases} \phi_1^\alpha\alpha, \ \phi_2^\alpha\alpha, \ \cdots, \ \phi_{N/2}^\alpha\alpha \\ \phi_1^\beta\beta, \ \phi_2^\beta\beta, \ \cdots, \ \phi_{N/2}^\beta\beta \end{cases} \tag{1.40}$$

In the restricted Hartree-Fock model (RHF), each spin orbital is associated with a sibling that shares the same spatial function but carries the opposite spin function

$$|\phi_1^\alpha\rangle = |\phi_1^\beta\rangle \tag{1.41}$$

$$\vdots$$

$$|\phi_{N/2}^\alpha\rangle = |\phi_{N/2}^\beta\rangle \tag{1.42}$$

which results in the orthonormality between the α and β spatial functions.

$$\langle \phi_i^\alpha | \phi_j^\beta \rangle = \delta_{ij} \tag{1.43}$$
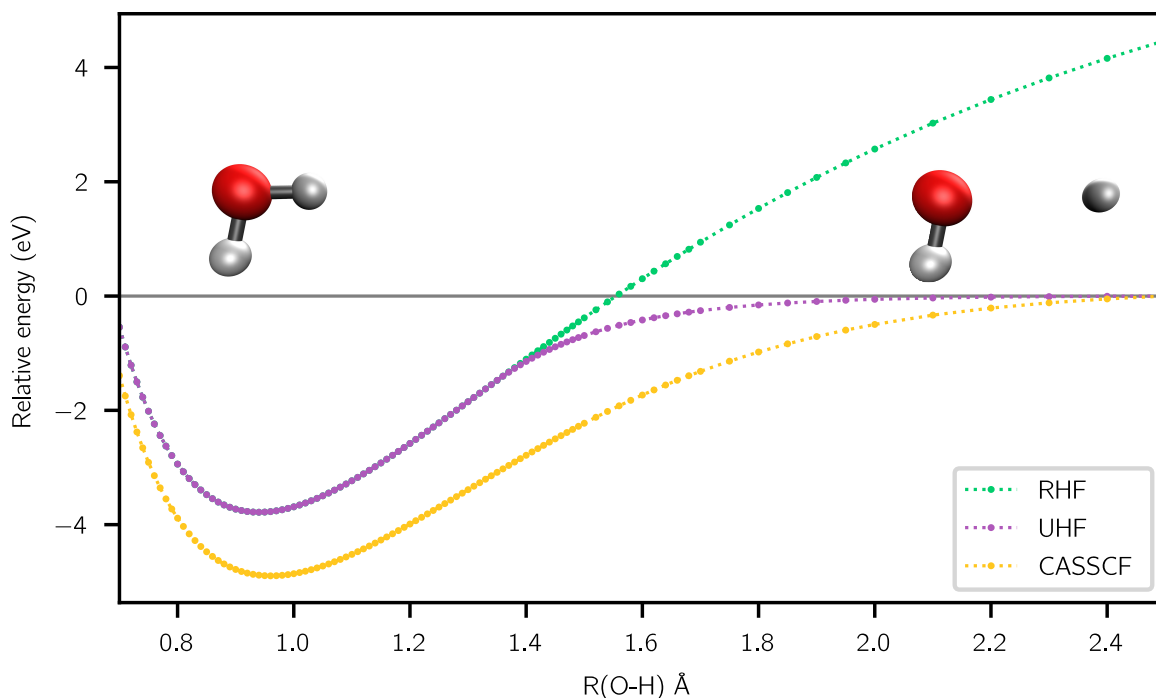
Figure 1.2: Comparison of the O-H bond dissociation energy curve for water with RHF and UHF models against exact full-valence results from CASSCF(8, 6) in the pcseg-2 basis.

Since each electron is paired up with another one in the RHF model it is defined exclusively for systems with an even number of electrons, and the expectation value of the S^2 operator can be derived to be 0.

$$\langle \Phi^{\mathrm{RHF}}| S^2 |\Phi^{\mathrm{RHF}}\rangle = 0 \tag{1.44}$$

RHF thus satisfies important property of a true electronic wave function - it carries an allowed spin multiplicity (Subsection 1.1.3).

As chemists are well aware of, molecules with an even number of paired-up electrons are a common scenario in organic chemistry under standard conditions and even a variety of transition metal complexes. As a result, RHF and the methods that elaborate upon it enjoy a broad range of applicability. [25] However, RHF fails under certain circumstances. Consider, for example, elongating one of the O−H bonds in the water molecule to dissociation. It is established that the resulting fragments should each carry one electron (homolytic dissociation) and the RHF criteria of forcing all electrons to occupy molecular orbitals in pairs fails to qualitatively describe this situation. Within UHF, the conditions in Eqs. 1.41 -

1.42 are not enforced, which allows independent spatial functions within the the α and β sets of orbitals and makes the model suitable to describe the electrons localized in each of the dissociated fragments, OH$^\bullet$ and H$^\bullet$. [24, 26] As is evident from Figure 1.2, the approximate dissociation energy provided by UHF is vast improvement over the catastrophic failure of that provided by RHF.

As Martin Head-Gordon (my advisor) sometimes says: "there is no free lunch in quantum chemistry". The UHF model suffers from two unattractive features as pay for its improved energetics. First, the fact that each spin space has independent spatial functions results in the potential non-orthogonality between the two sets, a phenomena known as spin polarization. Note, however that the orthogonality between the α and β spin functions ensures the orthogonality of the spin orbitals themselves. Second, a UHF wave function potentially fails to satisfy one of the conditions of the exact wave function - carrying a specific spin. In the simplest possible case of homolytic dissociation, such as in H$_2$, giving both unpaired electrons either an α spin or a β spin results in a valid S^2 expectation value characteristic of a triplet state.

$$\langle \uparrow\uparrow | S^2 | \uparrow\uparrow \rangle = 2 \tag{1.45}$$
$$\langle \downarrow\downarrow | S^2 | \downarrow\downarrow \rangle = 2 \tag{1.46}$$

However, when the electron in one fragment is given α spin and the other a β spin the resulting UHF wave function carries an invalid S^2 expectation value.

$$\langle \uparrow\downarrow | S^2 | \uparrow\downarrow \rangle = 1 \tag{1.47}$$
$$\langle \downarrow\uparrow | S^2 | \downarrow\uparrow \rangle = 1 \tag{1.48}$$

S^2 in this case is half-way in between a singlet ($S^2 = 0$) and a triplet ($S^2 = 2$). Constructing a wave function that satisfies spin symmetry (i.e. is a spin eigenfunction) requires a linear superposition of the two configurations above, thus placing it beyond the HF model. If the sole purpose is to describe energetics with the Hartree-Fock model for a scenario where electrons must unpair, UHF performs better. However, spin-symmetry breaking may have detrimental consequences for more sophisticated calculations relying on the Hartree-Fock reference wave function.

I wrap up the discussion on RHF and UHF with a notational convenience which will be useful in the following sections. I will refer to a spin orbital simply by its molecular orbital index with a bar above to denote it carries a β spin function (and an α without it). For example:

$$|i\rangle = |\phi_i\rangle |\alpha\rangle \tag{1.49}$$
$$|\bar{j}\rangle = |\phi_j\rangle |\beta\rangle \tag{1.50}$$

Roothaan-Hall equations

The final step before arriving at a procedure amenable for computation rests in establishing a form for the molecular orbitals. The most common approach for calculations on molecular

systems (as opposed to periodic systems) consists of expressing the molecular orbitals as linear **c**ombinations of **a**atomic **o**rbitals $\omega(\mathbf{r})$, or the LCAO approach. [27, 28] In turn, the most common form of AOs consists of Gaussian radial distributions of different spread centered about each atom, and in products with Cartesian functions to create angular nodes; [29] the collection of AOs span what is referred to as the one-electron basis. Ideally, an AO basis would be complete and thus able to describe an arbitrary one-electron function but in practice quantum chemists must truncate it to express the most relevant parts of the one-electron space in the most compact manner. [30] Per the LCAO technique, an MO ϕ_i^σ is described in terms of an AO basis $\{\,\omega_\mu\,\}$ as

$$\phi_i^\sigma = \sum_\mu C_{\mu i}^\sigma \omega_\mu \tag{1.51}$$

Expressing the Hartree-Fock equation in matrix form, while expressing the spatial form of the SOs in the AO basis leads to the Roothaan-Hall equations. [31, 32]

$$\mathbf{FC} = \mathbf{SCE} \tag{1.52}$$

The matrix $\mathbf{C}$ collects the MO coefficients $C_{\mu i}^\sigma$, and can thus be thought of as a change of basis matrix. The matrix $\mathbf{S}$ serves as a metric accounting for the possible non-orthogonality of the AO basis.

$$S_{\mu\nu} = \langle \omega_\mu | \omega_\nu \rangle \tag{1.53}$$

The Hartree-Fock density matrix in the AO basis

$$\mathbf{P} = \mathbf{C}_{\mathrm{occ}} \mathbf{C}_{\mathrm{occ}}^{\mathrm{T}} \tag{1.54}$$

depends only on the MO coefficients of the occupied SOs. The Fock matrix can be expressed in terms of $\mathbf{P}$ instead of $\mathbf{C}$. In the RHF model, for example, where the density associated with the α and β SOs is equal, the Fock matrix elements can be written in the AO basis as

$$\mathbf{F}_{\mu\nu} = \mathbf{h}_{\mu\nu} + \sum_{\lambda\sigma} \mathbf{P}_{\lambda\sigma} \left(\langle \mu\sigma | \nu\lambda \rangle - \frac{1}{2} \langle \mu\sigma | \lambda\nu \rangle \right) \tag{1.55}$$

where I've introduced the abbreviation for the two-electron integrals in Dirac or physisist notation.

$$\langle \mu\sigma | \nu\lambda \rangle = \int d\mathbf{r}_1 d\mathbf{r}_2 \frac{\omega_\mu^*(\mathbf{r}_1)\omega_\sigma^*(\mathbf{r}_2)\omega_\nu(\mathbf{r}_1)\omega_\lambda(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} \tag{1.56}$$

Crucially, unitary rotations among the occupied orbitals themselves and among virtual orbitals themselves

$$\mathbf{U}^\dagger \mathbf{U} = \mathbf{1} \tag{1.57}$$

$$\mathbf{U} = \begin{pmatrix} \mathbf{U_{oo}} & \mathbf{0} \\ \mathbf{0} & \mathbf{U_{vv}} \end{pmatrix} \tag{1.58}$$

which define a a new set of MO coefficients

$$\mathbf{C}' = \mathbf{C}\mathbf{U} \tag{1.59}$$

result in no changes to the Fock matrix Eq. 1.55. This makes it clear that all the relevant information is encoded in the space spanned by the occupied SOs, and not in the exact definitions of the orbitals themselves. An important note here is that, since the Fock operator depends on its eigenfunctions through the density matrix, the Roothaan-Hall equations are a pseudo-eigenvalue problem.

Self-consistent field solvers

The Roothaan-Hall equations are usually solved by proposing a guess for the MOs to construct a Fock operator, solving the pseudo-eigenvalue problem, harvesting an updated set of MOs and iterated until self-consistency: when Fock operator constructed from a set of MOs yields the same MOs as eigenvectors. While they were proposed decades ago, the Roothaan-Hall or self-consistent field (SCF) equations remain central to modern electronic structure and therefore quantum chemists have devised numerous ways to provide them an accurate initial guess [33] and proposed specialized algorithms to speed up their convergence. Below I present two such algorithms that I've used during my graduate work.

The direct inversion of the iterative subspace (DIIS) procedure involves using the information encoded in the convergence of the density after several SCF iterations to propose an improved guess on a subsequent iteration. [34–36] An error vector $\mathbf{e}$ is defined for the previous m SCF iterations, and an error vector for iteration $(m + 1)$ is approximated as

$$\mathbf{e}_{m+1} = \sum_m c_m \mathbf{e}_m \tag{1.60}$$

DIIS seeks to minimize $\mathbf{e}_m$ and use the resulting coefficients to construct an updated guess of the Fock matrix or the density as a linear combination of those on previous iterations. While this algorithm has found ample use in electronic structure for the purposes of accelerating the SCF convergence and others, organic radicals and transition metal complexes where several stationary points in orbital space exist near the desired minimum often pose a challenge.

The geometric direct minimization (GDM) algorithm provides a more robust alternative to DIIS. [37] It relies on the fact that, per the discussion on the previous subsection, only occupied-virtual orbital rotations modify the SCF energy

$$\mathbf{V} = \begin{pmatrix} \mathbf{0} & \mathbf{U_{ov}} \\ \mathbf{U_{vo}} & \mathbf{0} \end{pmatrix} \tag{1.61}$$

Moving along the direction of steepest-descent of the energy

$$\mathbf{G} = \frac{dE}{d\mathbf{V}} \tag{1.62}$$

dressed with approximate Hessian information, provides an SCF solver relying on a proper minimization algorithm.

Correlation energy

At this point, an important remark is that since an infinite basis set is required to span the whole space of arbitrary one-electron functions, there is correspondingly an infinite number of spin orbitals. However, only N spin orbitals are required to build the Hartree-Fock wave function; usually, the N spin orbitals with the lowest orbital energies are employed for that purpose. Of what use are then the remaining?

Recall that, as established by the variational principle (Eq. 1.31), the Hartree-Fock energy provides an upper bound to the exact electronic energy of the ground state; the remainder is given the name correlation energy.

$$\mathcal{E}_0 = E^{H.F.} + E_{corr.} \tag{1.63}$$

Post-Hartree-Fock methods often rely on the unoccupied spin orbitals to construct improved wave functions that can capture the correlation energy. Out of all, configuration interaction provides the simplest framework to illustrate this idea. [10, 38]

1.2.2 Configuration interaction

Hartree-Fock partitioning and the many-electron basis functions

Imagine, once again, partitioning the spin orbitals as those that are employed to construct the Hartree-Fock wave function (occupieds) and those that are not (unoccupied or virtual orbitals.) Traditionally, the indexes $i, j, k, ...$ are reserved for the N_{occ} occupied SOs; the indexes $a, b, c...$ to the N_{vir} virtual SOs; and the indexes $p, q, r, ...$ to arbitrary (occupied or virtual) SOs in the one-electron basis, N. Within the context of second-quantization, elaborated on in the upcoming coupled-cluster Subsection, the occupied-virtual partitioning defines a vacuum state and $|\Phi_0\rangle$ is also given the symbol $|\varnothing\rangle$ to remind us of it.

Excited configurations, or Slater determinants orthogonal to $|\varnothing\rangle$, can be constructed by replacing occupied spin orbitals with virtual ones. For book-keeping, a particular excited configuration is classified within an excitation rank. For example, a singly-excited configuration $|\Phi_i^a\rangle$ results when a single occupied spin orbital is swapped by a virtual spin orbital; a doubly-excited configuration $|\Phi_{ij}^{ab}\rangle$ results from swapping two spin orbitals; so forth and so on... The subsets of all singly-, doubly-, ... n-tuply-excited configurations are collected within the symbols

$$|S\rangle = \{ \ | \ \Phi_i^a\rangle \ | \ i \in \text{occ. and } a \in \text{vir.} \} \tag{1.64}$$

$$|D\rangle = \{ \ | \ \Phi_{ij}^{ab}\rangle \ | \ i, j \in \text{occ. and } a, b \in \text{vir.} \} \tag{1.65}$$

$$\vdots$$

Subject to the constraints of the one-electron basis, the set of all possible configurations forms a complete n-electron basis spanning the space of an arbitrary n-electron function, including the true (ground- and excited-state) electronic wave functions $|\Psi_k\rangle$.

$$\text{span}\{|\Psi_k\rangle\} = \text{span}\{|N\rangle\} \tag{1.66}$$

Full configuration-interaction

A natural way to express $|\Psi_0\rangle$ is as a simple linear combination of the HF reference function with all the excited configurations.

$$|\Psi_0\rangle = C_0 |\varnothing\rangle + \sum_{i,\,a} C_i^a |\Phi_i^a\rangle + \frac{1}{4} \sum_{i,\,j,\,a,\,b} C_{ij}^{ab} |\Phi_{ij}^{ab}\rangle + \cdots \tag{1.67}$$

The eigenvector with the lowest eigenvalue of the Hamiltonian in the complete n-electron basis,

$$
\mathbf{H} =
\begin{matrix}
 & |\varnothing\rangle & |S\rangle & |D\rangle & \cdots & |N\rangle \\
\langle\varnothing| & \begin{bmatrix} E^{H.F} & \mathbf{H}_{0S} & \mathbf{H}_{0D} & \cdots & \mathbf{H}_{0N} \\ & \mathbf{H}_{SS} & \mathbf{H}_{SD} & \cdots & \mathbf{H}_{SN} \\ & & \mathbf{H}_{DD} & \cdots & \mathbf{H}_{DN} \\ & & & \ddots & \vdots \\ & & & & \mathbf{H}_{NN} \end{bmatrix}
\end{matrix}
$$

defines the CI coefficients and the energies of the exact ground state wave function. While straightforward in concept, the aforementioned procedure - full CI - is severely limited in applicability. The number of all possible configurations for an n-electron system, described with a one-electron basis of size N is

$$\binom{N}{n} \tag{1.68}$$

Consider water, with 10 electrons, as an illustrative example. Nowadays, the size of the one-electron basis for a routine molecular calculation on water could be on the order of $N \approx 100$. The number of all determinants in this case is on the order of $\approx 10^{13}$. A great deal of work has thus gone into finding and testing a variety of different ways to consider only the most relevant configurations. Following the ethos of the field, scientists have sought approaches to approximations to the FCI problem, or ways to reduce its dimensionality, while retaining the most relevant configurations. In the final four subsections we mention key approaches or notions that are relevant to CI and other methods in electronic structure, as well as the novel work introduced in this thesis.

Brillouin's theorem

Brillouin provided us the important realization that the Hamiltonian matrix element connecting the ground state and any singly-excited configuration

$$\langle\varnothing| H |\Phi_i^a\rangle = f_{ia} \tag{1.69}$$

must be zero when the Hartree-Fock equations are satisfied since it is equal to an off-diagonal matrix element of the (diagonal) Fock operator; the whole ground-singles block of the Hamiltonian is then zero, $\mathbf{H}_{\varnothing S} = 0$. [39]

Truncated configuration-interaction

Another fact, crucial to electronic structure, arises from the Hamiltonian being a two-body operator: it depends at most on the coordinates of two electrons, via V_{ee} (Eq 1.5). As a result, the Hamiltonian can connect configurations that differ by at most two spin orbitals, so that $\mathbf{H}_{\varnothing N}$, where $N > 2$, must be zero. The only non-zero block of the Hamiltonian connecting the ground state is thus $\mathbf{H}_{\varnothing D}$, which alludes to the dominant role of doubly-excited configurations in much of electronic structure. Singly, triply, ... N-tuply excited configurations contribute passively to the ground state electronic wave function and energy via either direct or indirect configuration interaction with doubly-excited configurations. While a simple approach to approximating the correlation energy could be to consider the Hamiltonian within the sub-space of single and double excitations (CISD), truncated CI formalisms suffer from lack of so-called size-consistency. For the purposes of this work, its sufficient to state that an anzats that fails to satisfy size consistency suffers from an approximate correlation energy that deteriorates with the size of the system being treated, rendering it unusable for chemical purposes.

Configuration state functions

As illustrated by bond dissociation within the UHF model (Subsection 1.2.1), determinants where two or more electrons are fully unpaired are, in general, not eigenfunctions of the S^2 operator. Excited-configurations feature the same behavior as UHF determinants for the same reasons. A useful approach is to construct explicit spin eigenfunctions out of excited configurations to target the multiplicity of the desired state (configuration state functions, or CSFs) and thus reduce the dimensionality of the CI problem by excluding CSFs of other multiplicities. As the exact non-relativistic wave functions belong to a specific multiplicity, we can rest assured that no harm is done in focusing exclusively within the target spin state.

A variety of procedures for generating CSFs out of a set of configurations is provided by Paunz. [40] My favorite involves diagonalizing the S^2 operator within the subspace of relevant configurations. For example, two $M_S = 0$ configurations with open-shells in the spatial orbitals (p, q) are possible and there is correspondingly two CSFs: one singlet and one triplet.

$$|^1\Phi_p^q\rangle = (2)^{-1/2}\left(|\Phi_p^q\rangle + |\Phi_{\bar{p}}^{\bar{q}}\rangle\right) \tag{1.70}$$

$$|^3\Phi_p^q\rangle = (2)^{-1/2}\left(|\Phi_p^q\rangle - |\Phi_{\bar{p}}^{\bar{q}}\rangle\right) \tag{1.71}$$

CSFs will make a key appearance in Chapter 3 of this thesis, for the purposes of describing core excited states atop valence excited states.

Multi-configurational character and strong correlation

I conclude the discussion on CI theory by introducing two concepts related to the appropriateness of a single Slater determinant to approximate a given electronic state: the multi-configurational character of the wave function and the degree of strong correlation present in

the system. While multi-configurational character and strong correlation some times come hand-in-hand, the former is more concerned with the nature of the wave function and the latter with the energy afforded by it.

A system is said to be multi-configurational, as opposed to single-reference, when more than a single Slater determinant is required to qualitatively represent the true wave function, which could arise for two reasons. First, the exact wave function must respect the symmetries of the Hamiltonian. For example, as explained in the Subsection on CSFs and Section 1.1.3, multiple configurations are often required to construct a qualitative approximation to the wave function that is an eigenfunction of the spin-squared operator; this is the norm for singly-excited singlet states of closed-shell systems. Another example symmetry is the spatial one imposed by the nuclear framework which, for example, results in multi-configurational character for diatomics with Π term-symbols. Second, leaving symmetries aside, it is possible that multiple Slater determinants are close in energy and, as a result, make significant and comparable contributions to the CI expansion of the wave function. Such a situation arises, for example, in the ground state of the beryllium atom, [41, 42] some transition metal complexes, [25, 43] and organic molecules with multiple bonds stretched beyond their natural length. [44]

The correlation energy (Eq 1.63) is often classified as either dynamic or static. When a single Slater determinant is an appropriate zeroth-order wave function, as it plays out in a vast number of stable molecules, the contribution of any individual excited configuration is close to negligible. The collective contribution to the energy by the excited configurations - the dynamic correlation - may nonetheless be considerable and is of prime importance for chemistry. Symmetry-breaking excluded, when a single Slater determinant fails to capture the qualitative zeroth-order description of the wave function and multi-configurational character arises, the correlation energy is termed strong correlation. While dynamic and static correlation are well defined conceptually, the quantitative measures and thresholds used to classify the correlation energy as one or the other remain debated to this day. [25, 45–48]

In this thesis, I focus exclusively on molecular systems with ground states of closed-shell, single reference character, featuring no strong correlation. A variety of electronic structure methods specialized to this scenario have been developed since the beginnings of quantum chemistry. The following Sections review the ones relevant to this work: Møller-Plesset perturbation theory, coupled-cluster theory, and density functionl theory.

1.2.3 Møller-Plesset perturbation theory

Hamiltonian partitioning

An alternative approach to approximating the correlation energy comes from Rayleigh-Schrödinger (RS) perturbation theory (PT), wherein Møller and Plesset (MP) imagined the partitioning of the electronic Hamiltonian that bears their name. [49] In MP-PT, the Fock operator serves as the unperturbed, zeroth-order Hamiltonian, and a fluctuation potential

as the perturbation:

$$H = f + \lambda V_c \tag{1.72}$$
$$V_c = H - f \tag{1.73}$$

Since the perturbation parameter λ equals 1, it serves purely as a formal ordering parameter. Expanding the exact energy and wave functions to increasing orders in λ provides a central equation to derive the corrections to different order in PT by collecting terms of equal order and left-projecting with different configurations.

$$(f + \lambda V_c) \left| \Psi_0 \right\rangle = \mathcal{E}_0 \left| \Psi_0 \right\rangle \tag{1.74}$$
$$\left| \Psi_0 \right\rangle = \left| \varnothing \right\rangle + \lambda \left| \Phi^{(1)} \right\rangle + \lambda^2 \left| \Phi^{(2)} \right\rangle + \cdots \tag{1.75}$$
$$\mathcal{E}_0 = E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + \cdots \tag{1.76}$$

Note that in the expansion above, if the reference configuration $\left| \varnothing \right\rangle$ is taken to be normalized (Eq. 1.29) and its weight in the PT expansion equal to 1, then the expansion of the exact wave function in Eq. 1.75 is said to be intermediately-normalized. This means that its norm is greater than one but it can be re-normalized. The Hartree-Fock configuration, along with the rest of the n-electron basis introduced in Section 1.2.2, are eigenfunctions of the unperturbed Hamiltonian

$$f \left| \varnothing \right\rangle = E^{(0)} \left| \varnothing \right\rangle \tag{1.77}$$
$$f \left| \Phi_i^a \right\rangle = (E^{(0)} + \Delta_i^a) \left| \Phi_i^a \right\rangle \tag{1.78}$$
$$f \left| \Phi_{ij}^{ab} \right\rangle = (E^{(0)} + \Delta_{ij}^{ab}) \left| \Phi_{ij}^{ab} \right\rangle \tag{1.79}$$
$$\vdots$$

where $E^{(0)}$ is the sum of occupied orbital energies, and Δ correspond to orbital energy differences (Eq. 1.33)

$$E^{(0)} = \sum_i \epsilon_i \tag{1.80}$$
$$\Delta_i^a = \epsilon_a - \epsilon_i \tag{1.81}$$
$$\Delta_{ij}^{ab} = \epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j \tag{1.82}$$
$$\vdots$$

First- and second-order corrections to the energy

Truncating the PT at different orders (n = 1, 2 ...) defines the different MPn models. The contributions to the energy are obtained by left-projecting Eq 1.74 (collected to different

orders in λ) with the zeroth order wave function $\langle\varnothing|$.

$$E^{(0)} = \langle\varnothing| \, f \, |\varnothing\rangle \tag{1.83}$$

$$E^{(1)} = \langle\varnothing| \, V_c \, |\varnothing\rangle \tag{1.84}$$

$$E^{(2)} = \langle\varnothing| \, V_c \, |\Phi^{(1)}\rangle \tag{1.85}$$

$$E^{(3)} = \langle\varnothing| \, V_c \, |\Phi^{(2)}\rangle \tag{1.86}$$

$$\vdots$$

Related to Wigner's (2n + 1) rule for non-degenerate PT, the contribution to the energy $E^{(n)}$ requires the $|\Phi^{(n-1)}\rangle$ correction to the wave function. In this thesis we are concerned with corrections to the energy up to second order so we work our way up. The zeroth order correction to the energy is defined in Eq 1.80. Adding on the first-order correction to the energy, Eq. 1.84, to the zeroth order correction gives the HF energy.

$$E^{(0)} + E^{(1)} = \langle\varnothing| \, f + V_c \, |\varnothing\rangle \tag{1.87}$$

$$= \langle\varnothing| \, H \, |\varnothing\rangle \tag{1.88}$$

$$= E^{\text{H.F.}} \tag{1.89}$$

The first order correction to the wave function can be expressed in the basis of n-electron configurations as they form a complete basis (Eq 1.66)

$$|\Phi^{(1)}\rangle = \sum_N |N\rangle \, \langle N|\Phi^{(1)}\rangle \tag{1.90}$$

$$= \sum_N c_N^{(1)} \, |N\rangle \tag{1.91}$$

Expanding Eq. 1.74 in the perturbation parameter and collecting terms to first order defines the first-order wave function. Making use of the explicit form of the first-order correction to the energy (Eq. 1.84)

$$(E^{(0)} - f) \, |\Phi^{(1)}\rangle = (V - E^{(1)}) \, |\varnothing\rangle \tag{1.92}$$

$$(E^{(0)} - f) \, |\Phi^{(1)}\rangle = (V - \langle\varnothing| \, V \, |\varnothing\rangle) \, |\varnothing\rangle \tag{1.93}$$

Inserting Eq. 1.91 into Eq. 1.93 results in

$$\sum_N c_N^{(1)} (E^{(0)} - f) \, |N\rangle = (V - \langle\varnothing| \, V \, |\varnothing\rangle) \, |\varnothing\rangle \tag{1.94}$$

Left-projection with the different n-electron basis functions (Eqs. 1.64 and 1.65) and relying on the complex conjugate of Eqs. 1.77 - 1.79 provides explicit expressions for the coefficients

$c_N^{(1)}$.

$$c_\varnothing^{(1)} = 0 \tag{1.95}$$

$$c_S^{(1)} = \frac{\langle S | V | \varnothing \rangle}{\Delta_S} \tag{1.96}$$

$$c_D^{(1)} = \frac{\langle D | V | \varnothing \rangle}{\Delta_D} \tag{1.97}$$

Since the correlation potential is a two-body operator, like the Hamiltonian, excitations beyond the rank of triples will result in zero for the left-hand side of Eq. 1.94. If the Hartree-Fock equations are satisfied and Brillouin's theorem (Eq. 1.69) holds, then Eq. 1.96 is zero and only doubly-excited configurations contribute to the first order correction to the wave function. Making use of the aforementioned discussion, and by Eqs. 1.85 and 1.65, the usual form of the second-order correction to the electronic energy (MP2) emerges

$$E^{(2)} = - \sum_D \frac{\langle \varnothing | V_c | D \rangle \langle D | V | \varnothing \rangle}{\Delta_D} \tag{1.98}$$

$$= - \frac{1}{4} \sum_{ijab} \frac{| \langle \varnothing | V_c | \Phi_{ij}^{ab} \rangle |^2}{\Delta_{ij}^{ab}} \tag{1.99}$$

Note a factor of $\frac{1}{4}$ has been introduced to account for the double-counting in the unrestricted sums over (i, j) and (a, b). In general, for every unrestricted sum over n indexes belonging to the occupied subspace contributes a factor of $\frac{1}{2^{n-1}}$ to account for double counting. The same holds for unrestricted sums over n indexes belonging to the virtual subspace.

While not variational, MP-PT is size-consistent to all orders and it is well-suited for single-reference, dynamically-correlated molecular systems. For some such systems, Cremer and He showed how, even at second order, MP-PT captures a large fraction of the correlation energy and converges smoothly and quickly upon increasing to higher orders. [50] The same article, however, also illustrates cases where MP-PT converges erratically if at all. These are a manifestation of the fact that MP-PT is sensitive the the quality of the zeroth-order wave function which causes it to fail on the early onset of strong-correlation. In an effort to turn this sensitivity into an advantage, and make MP-based methods more robust to calculations on radicals and transition states, research in the last decade has explored the use of improved reference wave functions from explicit orbital-optimization in the presence of the MP correlation energy or from density functional theory. [51–53] Investigations of MP-based methods by empirically scaling the same-spin and opposite-spin contributions to the correlation energy to improve its performance for challenge cases, applying regularizations to the energy to tame divergences, and developing new algorithms to improve the computational efficiency speak to the prominent role of this theory to date. [52, 54–59] Regardless, some tasks still belong to the more accurate, robust, and expensive cousin of MP-PT: coupled-cluster theory.

1.2.4   Coupled-cluster theory

Notation of second-quantization

To assist in the discussion of coupled-cluster theory, I introduce the language of second quantization within the context of the electronic problem. [60] To generate excited configurations, second quantization defines an "annihilation" operator $\hat{a}_p$ which serves to delete the row of a Slater determinant associated with spin orbital χ_p (once moved all the way to the first index, while keeping track of changes in the sign of the determinant), and a "creation" operator $\hat{a}_q^\dagger$ which inserts the spin orbital χ_q at the first index. Since the wave functions in this thesis are concerned exclusively with electrons and not any other kind of elementary particle I omit the $\hat{a}$ for individual annihilation and creation operators and refer to them exclusively by their spin orbital index. For example, a particular singly- or doubly-excited configuration can be generated by action of the appropriate operators on the vacuum state

$$a^\dagger i \left|\varnothing\right\rangle = \left|\Phi_i^a\right\rangle \tag{1.100}$$

$$a^\dagger b^\dagger j i \left|\varnothing\right\rangle = \left|\Phi_{ij}^{ab}\right\rangle \tag{1.101}$$

$$\vdots$$

Furthermore, let us define an excitation operator $\hat{\alpha}_{ijk\cdots}^{abc\cdots}$ that corresponds to a string, or product, of annihilation operators $(\cdots kji)$ to the right of creation operators $(a^\dagger b^\dagger c^\dagger \cdots)$. For example,

$$\hat{\alpha}_i^a = a^\dagger i \tag{1.102}$$

$$\hat{\alpha}_{ij}^{ab} = a^\dagger b^\dagger i j \tag{1.103}$$

$$\vdots$$

To refer to an arbitrary n-electron excitation operator, I use the symbol $\hat{\alpha}_n$.

Coupled-cluster ansatz

Coupled-cluster (CC) theory proposes a systematically improvable ansatz for the ground state electronic wave function. [61, 62] It takes the form

$$\left|\Psi_0\right\rangle = e^{\hat{T}} \left|\varnothing\right\rangle \tag{1.104}$$

where the cluster operator generates all possible excitations and tags them along with a coefficient

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \cdots + \hat{T}_N \tag{1.105}$$

$$\hat{T}_1 = \sum_{ia} t_i^a \hat{\alpha}_i^a \tag{1.106}$$

$$\hat{T}_2 = (2)^{-2} \sum_{ijab} t_{ij}^{ab} \hat{\alpha}_{ij}^{ab} \tag{1.107}$$

$$\vdots$$

$$\hat{T}_n = (n!)^{-2} \sum_{n} t_n \hat{\alpha}_n \tag{1.108}$$

The coupled-cluster ansatz is exact and maps one-to-one with the FCI wave function; it also inherits its steeply-scaling complexity. For practical purposes, the cluster operator is truncated to some excitation rank to provide an approximate ground state wave function. For example the coupled-cluster singles and doubles (CCSD) includes single and double excitations, resulting in a wave function

$$|\Phi^{\mathrm{CCSD}}\rangle = e^{(\hat{T}_1 + \hat{T}_2)} |\varnothing\rangle \tag{1.109}$$

The variational principle is incompatible with CC theory for routine use since evaluating the expectation value of the Hamiltonian with the CC wave function requires exponential computations even when truncated to a subspace. [61] Instead, the CC theory amenable for practical computation ("projected" CC) relies on the fact that a similarity transformation leaves the eigenvalues of a Hermitian operator, such as the electronic Hamiltonian, unchanged.

$$\bar{H} = e^{-T} H e^{T} \tag{1.110}$$

The similarity transformation affects the eigenvectors of the operator as well

$$|\bar{\Phi}^{\mathrm{CCSD}}\rangle = e^{-\hat{T}} |\Phi^{\mathrm{CCSD}}\rangle \tag{1.111}$$

$$= e^{-\hat{T}} e^{\hat{T}} |\varnothing\rangle \tag{1.112}$$

$$= |\varnothing\rangle \tag{1.113}$$

The approximate energy within the non-Hermitian, non-variational CC model is then given by the expectation value of the similarity-transformed Hamiltonian with the vacuum state

$$E^{\mathrm{CC}} = \langle\varnothing| \bar{H} |\varnothing\rangle \tag{1.114}$$

The evaluation of the CC energy relies on the BCH theorem, which ends up truncating at the fourth-degree nested commutator. Enough equations to define the coefficients of the CC

wave function (Eqs. 1.104 - 1.108), required to evaluate the CC energy, are generated by projection of the Schrödinger equation with the different excited configurations within the relevant space.

$$\langle \Phi_i^a | \bar{H} | \varnothing \rangle = 0 \tag{1.115}$$

$$\langle \Phi_{ij}^{ab} | \bar{H} | \varnothing \rangle = 0 \tag{1.116}$$

$$\vdots$$

Note that the resulting equations are complicated and non-linear, since the left-hand side depends on the amplitudes themselves. Like the SCF procedure, the CC equations are solved iteratively by:

1. proposing a guess for the amplitudes. Usually, these are the first-order correction to the wave function from MP2 theory.

2. solving the equations for a new set of amplitudes.

3. iterating to self consistency. Acceleration procedures like DIIS are often employed. [63]

Compared with a version of CISD corrected for size-consistency, [64] CCSD provides similar correlation energies for single-reference systems. [65, 66] Part of the appeal of CC theory lies in that it provides size-consistent energy out the box at any level of truncation. The latter is illustrated by the exponential ansatz of the wave function, which allows for a separation of the cluster operators associated with two independent, non-interacting subsystems. [61] In addition to being size-consistent, CCSD is exact for 2-electron systems and thus a powerful generalization of HF theory which is size-consistent and exact for 1-electron systems.

1.2.5 Density functional theory

Density functional theory takes a fundamentally different approach than the methods mentioned previously. [67, 68] The very compact one-electron density for a n-electron system

$$\rho(\mathbf{r}) = N \int d\sigma_1 d\mathbf{x}_2 \cdots d\mathbf{x}_n |\Psi(\mathbf{x}_1, \cdots, \mathbf{x}_n)|^2 \tag{1.117}$$

takes the place of the mighty electronic wave function as the central quantity for evaluating the energy. Note that the ρ depends exclusively on the spatial coordinates of one electron by virtue of the integration over the rest of the electronic coordinates and the spin coordinate of the last electron. As a result, it is a vastly simpler object than the $3n$-dimnensional n-electron wave function.

Hohenberg-Kohn theorem and universal density functional theory

The external potential is the only piece of the Hamiltonian that distinguishes one system with n electrons from another. For our molecular Hamiltonian, this is simply the electrostatic potential between the electrons and the nuclei.

$$V_{eN} = -\sum_{i,\,A} \frac{Z_A}{\mathbf{r}_i - \mathbf{R}_A} \tag{1.118}$$

$$= \sum_i v_{ext}(\mathbf{r}_i) \tag{1.119}$$

Solving the Schrödinger equation and using the resulting wave function to evaluate the one-electron density establishes a mapping between the external potential and the one-electron density.

$$v_{ext} \longrightarrow \rho \tag{1.120}$$

In 1964, Hohenberg and Kohn established that this mapping can be inverted, up to a constant. [69]

$$\rho \longrightarrow v_{\text{ext.}} + C \tag{1.121}$$

Since the one-electron density can define the external potential, the n-electron wave function itself can be expressed as a functional of the former. The term universal density functional was coined for any system sharing the same number of electrons.

$$F[\rho] = \langle \Psi[\rho]|\, \hat{T} + \hat{V}_{ee}\, |\Psi[\rho]\rangle \tag{1.122}$$

Therefore, the total electronic energy of the ground state is itself a functional of the one-electron density.

$$E_0[\rho] = F[\rho] + \int d\mathbf{r}\rho(\mathbf{r})v_{\text{ext.}}(\mathbf{r}) \tag{1.123}$$

Finally, Hohenberg and Kohn showed that the exact electronic energy is obtained by minimizing the energy functional with respect to v-representable one-electron densities, which are those associated with a non-degenerate ground state associated with a local external potential, when in possession of the exact universal functional. [70]

$$\mathcal{E}_0 = \min_{\rho}\{E_0[\rho]\} \tag{1.124}$$

However, Levy found v-representability to be too strict a demand, and proved that the minimization can be carried instead over n-representable densities, which are those associated with antisymmetric, normalized n-electron wave functions. [71, 72] The exact form of the universal density functional remains unknown to this day, and early approximations to it proved unsuccessful for molecules. Nonetheless, research into the universal functional or "orbital-free" formalism of DFT remains active. The advent of the Kohn-Sham method, however, turned density functional theory into the successful electronic structure model that it is today.

Kohn-Sham density functional theory

Kohn and Sham came up with the idea of breaking the universal energy functional into several components: a kinetic energy component, a Coulomb term, and an exchange-correlation component. [73]

$$F[\rho] = T_s[\rho] + E_J + E_{xc}[\rho] \tag{1.125}$$

The kinetic energy piece is defined by minimizing the expectation value of a kinetic energy operator acting on a single-determinant wave function for a fictitious non-interacting system sharing the density of the true wave function, called the Kohn-Sham wave function.

$$T_s[\rho] = \min_{\rho}\{\langle\Phi^{\mathrm{KS}}|\,\hat{T}\,|\Phi^{\mathrm{KS}}\rangle\} \tag{1.126}$$

The Coulomb component is simply the classic electrostatic repulsion of a charge distribution given by ρ. In other words, it is equivalent to the sum of the standard Coulomb contributions to the electron-electron repulsion (Eq. 1.37).

$$E_J[\rho] = \frac{1}{2}\int d\mathbf{r}_1 d\mathbf{r}_2 \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} \tag{1.127}$$

The exchange-correlation functional can be written as

$$E_{xc}[\rho] = (\langle\Psi|\,\hat{T} + \hat{V}_{ee}\,|\Psi\rangle) - (T_s[\rho] - E_J[\rho]) \tag{1.128}$$

Importantly, this accounts for any differences in the kinetic energy of the true wave function and that of the non-interacting, fictitious Kohn-Sham wave function. In possession of the exact exchange-correlation functional, the Kohn-Sham formalism is exact. Needless to say, we do not know the form of the exact functional so a great deal of work has gone into finding approximations based on physical intuition, enforcing constraints on their form required by known constraints of the exact density, and fitting empirical parameters to reproduce either experimental data or high-level benchmark data coming from wave function theory. Before that, let us briefly introduce the form of the Kohn-Sham equations used in practice.

The Kohn-Sham equations

Operationally, the Kohn-Sham equations emerging from this formalism resemble the Hartree-Fock equations.

$$\hat{f}^{\mathrm{KS}}\,|\phi_p^{\mathrm{KS}}\rangle = \varepsilon_p\,|\phi_p^{\mathrm{KS}}\rangle \tag{1.129}$$

$$\tag{1.130}$$

The Kohn-Sham one-electron operator, defined in terms of the kinetic energy, the electron-nucleus potential, the Coulomb interaction, and the exchange-correlation potential, takes

the place of the Fock operator.

$$f^{KS}(\mathbf{r}) = h(\mathbf{r}) + v^{KS}(\mathbf{r}) \tag{1.131}$$

$$v^{KS}(\mathbf{r}) = v_{eN}(\mathbf{r}) + v_J + v_{\text{xc}}(\mathbf{r}) \tag{1.132}$$

$$v_{\text{xc}}(\mathbf{r}) = \frac{\delta E_{xc}[\rho]}{\delta \rho} \tag{1.133}$$

The Kohn-Sham orbitals $|\phi_i^{KS}\rangle$ describe the fictitious non-interacting electrons used to define the Kohn-Sham wave function Φ^{KS} and yielding, when the exact Kohn-Sham functional is used, the same density as the true wave function.

Approximations to the exchange correlation energy functional

The local density approximation (LDA) involves approximating the exchange-correlation per particle at each point in space as that of a uniform electron gas of the same density, for which the exact result is known. [73] Since molecules are basically inhomogeneous electron gases, as my advisor once put it, the next natural step involves including information about the gradient of the density, which gives rise to the gradient-expansion approximation (GEA) and its more successful cousin, the generalized gradient approximation (GGA). Hybrid and double hybrids add fractions of Hartree-Fock exchange and MP2 correlation, respectively, into the exchange correlation energy.

More sophisticated functionals, as measured by the information in the exchange-correlation functional, tend to perform better when evaluating energy differences between systems. [74] This led Perdew to establish the analogy between rungs of density functional approximations and the biblical ladder of Jacob. [75] However, this is not always the case and over-parametrization might result in functionals that perform spectacularly for the purpose they were designed but suffer from poor transferability across quantum chemical tasks. [75–78] Table 1.1 summarizes the different density functional rungs, along with a brief description of their functional form and their compute cost with respect to the size of the one-electron basis set N for a naive implementation.

Leaving questions of design aside, it is clear that density functional theory (DFT) is the standard tool of choice in electronic structure when it comes to real chemical applications. And this is with good reason when dealing with single-reference systems featuring no self-interaction error - the two great shortcomings of DFT - because the accuracy-cost ratio is better than that of wave function theory. [79]

1.3 Excited-state electronic structure

While HF, DFT, MP-PT, and CC provide the services enumerated in the opening paragraphs of Section 1.2 and many more, quantum chemistry is often required to go beyond the ground state. The electronic energy of the excited excited states

$$\mathcal{E}_k(\mathbf{R}) \tag{1.134}$$

Table 1.1: Jacob's ladder of density functional approximations.

Rung	Name	Form of E_{xc}	Compute cost
1	local density approximation (LDA)	$E_{xc}^{LDA}[\rho] = \int d\mathbf{r}\rho(\mathbf{r})\epsilon_{xc}^{unif.}(\rho)$ Approximates the correlation energy as each point with density ρ as that of the uniform electron gas with the same density, $\epsilon_{xc}^{unif.}(\rho)$	$O(N^3)$
2	generalized gradient approximation (GGA)	$E_{xc}^{GGA}[\rho] = \int d\mathbf{r}f(\rho, \nabla\rho)$ Incorporates information about the gradient of the density.	$O(N^3)$
3	meta-GGA	$E_{xc}^{mGGA}[\rho] = \int d\mathbf{r}f(\rho, \nabla\rho, \nabla^2\rho, \tau)$ Incorporates information about the Laplacian of the density and / or the kinetic energy density, $\tau = \frac{1}{2}\sum_i |\nabla\phi_i|^2$.	$O(N^4)$
4	hybrid-GGA	$E_{xc}^{hGGA}[\rho] = f(a \cdot E_x^{HF}, b \cdot E_{xc}^{GGA}, c \cdot E_{xc}^{LDA})$ Make linear combinations of HF exchange with GGA and LDA exchange correlation energies based on exact conditions and/or fit to experimental data or benchmark theory data.	$O(N^4)$
5	double hybrid	$E_{xc}^{DH}[\rho] = f(a \cdot E_x^{HF}, b \cdot E_c^{MP2}, c \cdot E_{xc}^{DFA})$ Make linear combinations of HF exchange and MP2 correlation with density func. approximations to the exchange and correlation.	$O(N^5)$

defines alternative PES landscapes where fascinating phenomena take place. For example, a molecule on a stable ground state might favor dissociation or bond rotation along some excited state PES. Matters are further complicated by the fact that excited PES often come close in energy or even touch, leading to non-adiabatic processes that break the Born-Oppenheimer approximation. This is all the realm of photochemistry.

The vertical excitation energy for the kth excited state corresponds to the difference from the ground state energy at a fixed nuclear geometry.

$$\Delta\mathcal{E}_{n\leftarrow 0}^{vert} = \mathcal{E}_k(\mathbf{R_x}) - \mathcal{E}_0(\mathbf{R_x}) \tag{1.135}$$

However, the minima for the excited state (if bound) may be different than that of the ground state. The difference in energy between the two electronic states, evaluated at their corresponding minima, is called the adiabatic excitation energy. This difference is illustrated
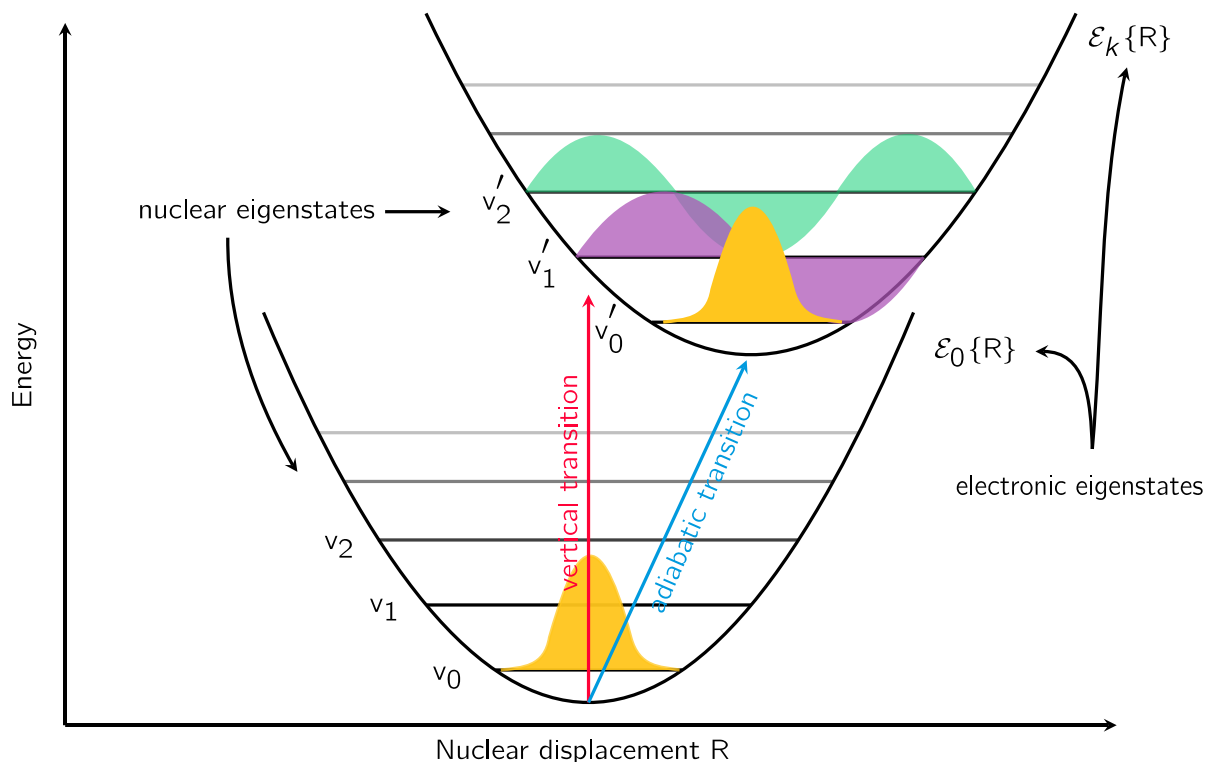
Figure 1.3: Illustration of the differences between vertical and adiabatic transitions.

in Figure 1.3.

$$\Delta \mathcal{E}_{n \leftarrow 0}^{adia.} = \mathcal{E}_k(\mathbf{R})|_{\mathbf{R}_k^{min.}} - \mathcal{E}_0(\mathbf{R})|_{\mathbf{R}_0^{min.}} \qquad (1.136)$$

Excitation energies provide energy locations for the transitions that could be observed in certain spectroscopies. Whether they are actually observed or not is dictated by the probability for the excitation taking place and the appropriate selection rules. In this regard, the quantum nature of the nuclei often plays an important role. The full molecular wave function, under the BO approximation, is truly a product of the electronic wave function and the nuclear wave function (Subsection 1.1.2).

$$|\Psi_{k, v}^{\text{mol.}}\rangle = |\Psi_k^{\text{el.}}\rangle |\Theta_v\rangle \qquad (1.137)$$

For one-photon absorption spectroscopies, Fermi's golden rule provides an approximation for the rate of change in time of the probability for a weak electric field, whose interaction with

the molecule is represented by $H' = H'_{\text{el.}} + H'_{\text{nuc.}}$, to induce a transition.

$$\Gamma_{k,\,v'\leftarrow0,\,v} = 2\pi \left\langle \Psi^{\text{mol.}}_{k,\,v'} \right| H' \left| \Psi^{\text{mol.}}_{0,\,v} \right\rangle \rho(E_{k,\,v'}) \tag{1.138}$$

$$= 2\pi \left[\left\langle \Theta_{v'} | \Theta_v \right\rangle \left\langle \Psi^{\text{el.}}_k \right| H'_{\text{el.}} \left| \Psi^{\text{el.}}_0 \right\rangle + \left\langle \Theta_{v'} \right| H'_{\text{nuc.}} \left| \Theta_v \right\rangle \left\langle \Psi^{\text{el.}}_k | \Psi^{\text{el.}}_0 \right\rangle \right] \rho(E_{k,\,v'}) \tag{1.139}$$

$$= 2\pi \left[\left\langle \Theta_{v'} | \Theta_v \right\rangle \left\langle \Psi^{\text{el.}}_k \right| H'_{\text{el.}} \left| \Psi^{\text{el.}}_0 \right\rangle \right] \rho(E_{k,\,v'}) \tag{1.140}$$

For the electronic contribution, the oscillator strength in terms of the transition dipole moment

$$f_{n\leftarrow0} = \frac{2}{3}\Delta\mathcal{E}_{n\leftarrow0} \times |\mu_{n0}|^2 \tag{1.141}$$

$$\mu_{n0} = \left\langle \Psi_n \right| \hat{\mu} \left| \Psi_0 \right\rangle \tag{1.142}$$

emerges from Fermi's golden rule by describing the electromagnetic field as a plane wave with a monochromatic frequency and under the dipole approximation, which holds when the wavelength is much larger than the extent over which the electrons are distributed. The oscillator strength thus provides a sort of measure for how likely an electronic transition is to occur. In regards to the nuclei, the Frank-Condon factor $\left\langle \Theta_{v'} | \Theta_v \right\rangle$ is simply the overlap between the nuclear wave functions in the ground and excited states and, under the right conditions (such as cold temperature), its effects can be observed: it confers electronic absorption spectra into bound excited states with a fine structure. Such a spectrum resembles a curve composed of jagged peaks, which reveal the quantum nature of the nuclei. Furthermore, the most intense of the peaks will often not be the lowest in energy, which is a consequence of the off-set between the minima in the two electronic surfaces. For wonderful examples of these phenomena, see Refs. [80–87]

Together, the excitation energy and the oscillator strengths afforded by an electronic structure method provide a way to simulate electronic absorption spectra of molecules. In this thesis we will work exclusively with vertical transitions and leave the fascinating nuclear phenomena aside. I begin the following subsections by providing a conceptual description of electronic excited states in chemical terms per molecular orbital theory, where electronic excited states are classified qualitatively as either promoting a valence or a core electron into a valence or a Rydberg state. Subsequently, I introduce some of the extensions to the aforementioned ground state methods and elaborate on their performance for describing core excited states.

1.3.1 Excited states within molecular orbital theory

Before delving into the mathematical machinery of electronic structure, let's define some key characteristics about excited states in terms of molecular orbital theory. As established previously, the electrons in the ground state inhabit orbitals in the occupied space. Since the electrons leave one (or multiple) of these orbitals during an electronic excitation, these depopulated orbitals are called hole levels within this context. An electron excited from a
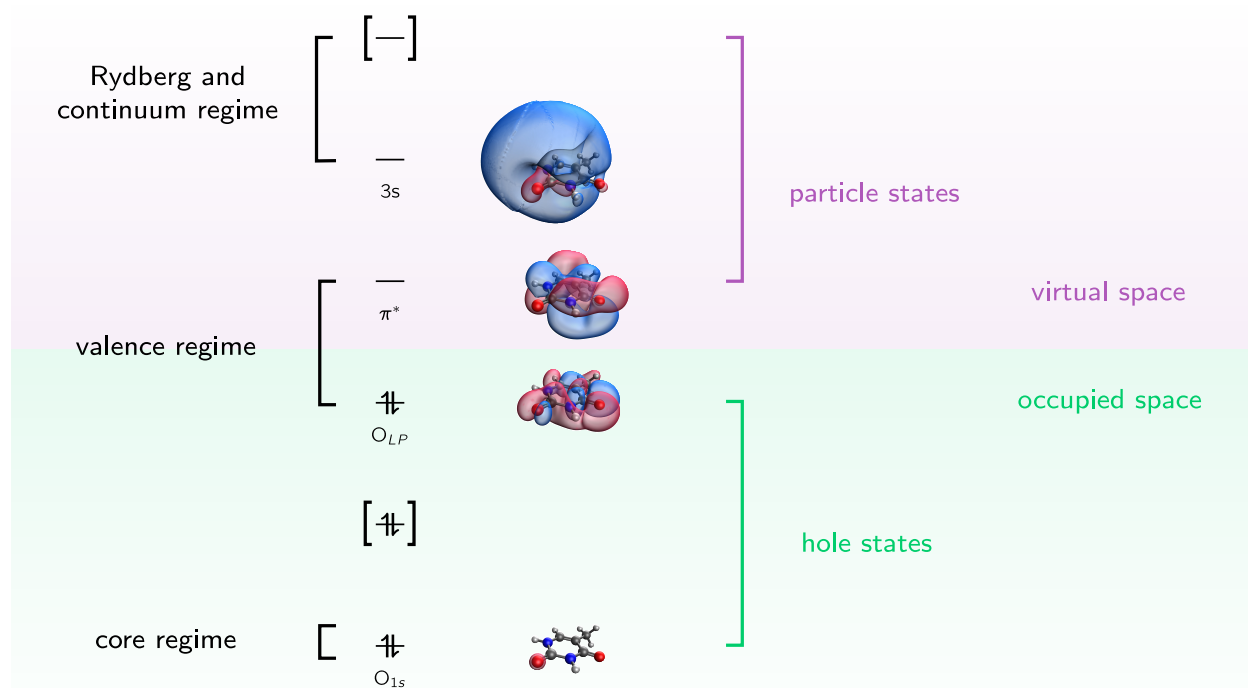
Figure 1.4: Illustration of the occupied-virtual orbital partitioning in thymine. Separating the different hole states as either core or valence, and the different particle states as either valence, Rydberg, or continuum.

hole level arrives somewhere: into one (or a superposition of several) particle level(s). The term "level" is used to specify that these are one-electron objects with which the eigenstates of the n-electron system can be conceptually characterized. Figure 1.4 uses thymine to illustrate the partitioning of orbitals into occupied and virtual, and their association with hole and particle levels. The characterization of the latter two comes below.

The hole state: core vs. valence

Hole states can be roughly classified into two categories: the core regime and the valence regime. The core electrons are the lowest-energy electrons which are strongly atomic in nature, meaning that they are predominantly confined to a region close to the nucleus. For the main group elements in a molecule, these are the orbitals arising from the 1s orbitals of its constituent atoms, such as one of the O_{1s} orbitals of thymine in Figure 1.4. Note that thymine has other core electrons associated with the other heavy atoms that are not

visualized in the Figure.

On the other extreme of the occupied orbital subspace is the valence regime, which arises from the fusion of valence AOs into delocalized MOs that spread over part or all of the molecule. In organic molecules, these are commonly π-bonding orbitals arising from the interaction between 2p orbitals perpendicular to bonds, σ-bonding orbitals arising from the interaction between 2s orbitals and 2p orbitals along bonds, and lone-pair electrons of heteroatoms. The highest-lying occupied orbital in thymine (Figure 1.4) is an example of the later.

The particle state: valence vs. Rydberg

The valence particle levels are often the low-energy anti-bonding analogues of the bonding-type orbitals in the occupied subspace. For example, the anti-bonding π^* orbitals in thymine arise as counterparts to the occupied π orbitals. Like their hole-level siblings, valence particle levels are often delocalized over a large part or all of the molecule.

Rydberg-type particle levels, on the other hand, come from 3s- and 3p-type empty orbitals in main-group elements and they are highly delocalized and diffuse. See, for example, the second particle level in thymine (Figure 1.4) seems to arise collectively from the 3s orbitals of several atoms in the molecule. As one climbs higher in energy along the virtual subspace, the manifold of particle levels becomes more dense, eventually converging to an ionization threshold beyond which an infinite number of continuum levels lies. Excited-state electronic structure methods most often focus on excited states below the ionization threshold.

Core level spectroscopy

For the first-row elements, the core regime corresponds to the electrons inhabiting the 1s orbitals. As the atomic number of the nuclei increases, the core regime moves down in energy due to the stronger electron-nuclei attraction between the core electrons and the nulei. While exciting a 1s electron takes roughly 280 - 300 eV for carbon atoms, it takes around 870 eV once you get to neon. Moving down to the third-row elements increases the required energy even more. Furthermore, for these elements, the 2s and 2p orbitals move from the valence into the core regime. To distinguish between the principal quantum number of the core electrons of interest, an alphabetic label is used. The convention is summarized in Table 1.2.

Table 1.2: Alphabetic label for core spectroscopies.

core-level principal quantum number	1	2	3	4	$\cdots$
Edge	K	L	M	N	$\cdots$

Moving beyond the K-edge requires also specifying the angular momentum of the electrons being probed. A subscript attached to the edge label is used to identify this. Figure

1.5 summarizes the terminology used to denote a core spectroscopy edge and provides the relevant energy ranges where these transitions take place.

E

M-edge

$M_{4,5}$ — — — — —
$3d_{x^2-y^2}$ $3d_{xy}$ $3d_{xz}$ $3d_{yz}$ $3d_{z^2}$

$M_{2,3}$ — — —
$3p_x$ $3p_x$ $3p_z$

M_1 —
$3s$

L-edge

$L_{2,3}$ — — —
$2p_x$ $2p_x$ $2p_z$

L_1 —
$2s$

	Na	Mg	Al	Si	P	S	Cl	Ar
$L_{2,3}$	21 to 41	41 to 61	63 to 83	89 to 109	122 to 142	155 to 175	191 to 211	237 to 257
L_1	53 to 73	79 to 99	108 to 128	139 to 159	179 to 199	219 to 239	270 to 290	310 to 330

K-edge

K —
$1s$

	C	N	O	F	Ne	Na	Mg	Al	Si	P	S	Cl	Ar
K	280 to 300	390 to 410	525 to 540	675 to 690	865 to 871	1060 to 1080	1300 to 1320	1550 to 1570	1830 to 1850	2135 to 2155	2462 to 2482	2812 to 2832	3192 to 3212

Figure 1.5: Energy ranges for different core edges for first and second group elements (eV).

As the energy ranges make clear, it takes a lot of energy to induce core transitions; the wavelengths of light required for K- and L-edge spectroscopies on the first and second row elements range from the extreme ultra-violet to the soft X-ray regimes. When the core electrons are excited into quasi-bound levels, below the core ionization thresholds, the spectroscopy is referred to as near-edge X-ray absorption spectroscopy (NEXAS, or simply XAS). Note that these are truly resonance states because, while they may be below the core ionization threshold, they are well above the valence ionization thresholds (on the order of 5 - 20 eV). If the lineshapes beyond the core ionization threshold are the subject of study instead, the technique is called extended X-ray absorption fine structure spectroscopy (EXAFS). X-ray photo-electron spectroscopy (XPS) is a complementary technique that aims to determine the core ionization threshold, or the core-binding energy.

The usually excellent energy separation of core regimes across elements is one of the reason for the appeal of core spectroscopies. Furthermore, they serve to probe the local environment around the atoms of a specific element in a molecule. For example, XAS and

EXAFS at the L- and M-edges of transition metal catalysts are often used to characterize the oxidation state of the metal and the metal-ligand distance under different conditions. [88] In the realm of fundamental physical chemistry, the ultrafast spectroscopy community has leveraged the element-specific sensitivity of core spectroscopy to uncover the signatures of excited states and nuclear motion. [89–94]

As valuable as core spectroscopies are, the progress of electronic structure to describe core transitions lags behind that of valence excited states due to a collection of challenges. The first is that they lie high in energy - well beyond the valence ionization threshold. The second is that, when a core electron is promoted, the remaining electrons suffer from a severe re-organization in response. In the following Subsections, I review some of the cornerstone excited state methods for bound or quasi-bound transitions (i.e. those below the ionization threshold of the hole level) and explore their performance for both core and valence excitations, while elaborating on the challenges of specific to core transitions and reviewing the strategies employed to circumvent them.

1.3.2 Configuration interaction singles

While truncated CI theory provides a poor framework to describe the ground state due to its failure to satisfy size-consistency, CI truncated to singly-excited configurations (CIS) serves as an excellent framework for qualitative approximations to some singly-excited electronic states. [95] As Brillouin's theorem (Eq. 1.69) makes clear, the Hartree-Fock ground-state energy and wave function remain unchanged after diagonalizing the Hamiltonian within the subspace of the reference and all singly-excited configurations. The remaining eigenvectors, higher in energy, must therefore correspond to approximate excited states and they each take the form

$$|\Phi_k^{\text{CIS}}\rangle = \sum_{ia}(b_i^a)_k |\Phi_i^a\rangle \tag{1.143}$$

When employing a RHF reference, it is easy to use CSFs instead of Slater determinants to target either singlet or triplet states and reduce the dimensionality of the CI matrix. The energies afforded by CIS for valence excited states are of a similar quality to the ground state energies provided by HF. For this reason, CIS and HF are both regarded as uncorrelated theories, the latter for the ground state and the former for the excited states.

While the dimensionality of the whole CIS matrix is (OV) $\times$ (OV), one is very often only interested in the first few eigenvalues, since these correspond to the valence excited states. To access only some of these, the ubiquitous iterative eigensolver of Davidson or some improvement of it is often employed leading to vast improvements in efficiency compared to the storage and diagonalization of the whole CIS matrix. [96, 97]

Natural transition orbitals

When expressing excited configurations with a large underlying AO basis in the canonical representation (i.e., the eigenfunctions of the Fock operator), it is common for the CIS expan-

sion to contain many numerically significant amplitudes. However, recall that the HF energy and density are invariant to occupied-occupied orbital rotations as well as virtual-virtual orbital rotations. Carrying out a singular value decomposition (SVD) on the amplitudes for each CIS state $\mathbf{b}_k$

$$\sigma_k = \mathbf{L}_k \cdot \mathbf{b}_k \cdot \mathbf{R}_k \tag{1.144}$$

provides a way carry out an occupied-occupied and virtual-virtual rotation of the MO basis to transform the CIS coefficients into a maximally compact representation. [98]

$$\mathbf{C}_k^{\text{NTO}} = \mathbf{C}_k \cdot \mathbf{R} \approx \begin{bmatrix} \vdots \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ \vdots \end{bmatrix}$$

The result is a condensed CIS expansion in the natural transition orbital (NTO) basis, where only a few terms (very often, to a good approximation, only one CSF) contribute significantly. As it turns out, the spatial extent of NTO orbitals resemble the picture provided by MO theory and are thus a useful tool to describe CIS excited states in a chemical way.

On the performance of CIS

The Quest data set [99, 100] provides high-level reference values for the following kinds of excitation energies:

- valence ($\pi \to \pi^*$, LP $\to \pi^*$) and Rydberg excitation energies, both singlet and triplet, for small and medium sized closed-shell molecules. [101, 102]

- excitation energies for radicals and organic molecules containing third-group elements. [103]

- charge transfer excitations in π-conjugated molecular systems. [104]

- excitations in bicyclic systems. [105]

On this data set, CIS achieves a root mean squared error (RMSE) of around 1.0 eV, with a mean-signed error (MSE) of 0.4 eV. [106] While this may not be enough for tasks requiring quantitative accuracy, it is certainly an excellent starting point.

Note, however, that this data set does not contain core excited states. The first challenge in applying CIS to core excitations is in that Davidson-based CI solvers must climb over the

large (infinite, formally) number of valence excited states before reaching the high-energy core regime. To that end, techniques such as the core-valence separation (CVS) scheme and the restricted-energy window (REW) approach were developed. [107, 108] With these techniques in place, it turns out that the performance of CIS for these types of excitations deteriorates by about one order of magnitude for a series of small molecules with only main-group elements. [109] Where does this discrepance in performance arise from?

From the perspective of orbitals, one can think of the effect in the valence electrons when a nuclear-charge-shielding core electron is excited or ionized altogether. The remaining electrons re-arrange and contract in response to the increased effective nuclear charge emanating from the nucleus previously hosting the core electron. This is a chemical way of saying that the excited state density differs significantly from that of the ground state, a phenomena that also occurs in charge-transfer and Rydberg excitations. CIS generates single excitations out of a Hartree-Fock reference optimized for the ground state so that, in a sense, it relies on poor ingredients to describe excited states. There will be much more on the issue of orbital relaxation in the following sections.

1.3.3 Time-dependent density functional theory

Like DFT, time-dependent density functional theory (TDDFT) is the most popular method for calculating approximate excited states. Let's review some of the fundamentals and explore the performance of TDDFT for valence and core excited states.

Runge-Gross theorem

In 1984, Runge and Gross generalized Hohenberg-Kohn theorem as well as the Kohn-Sham formalism to time-dependent potentials and densities. [110]

$$\rho(\mathbf{r}, t) \longleftrightarrow v_{ext}(\mathbf{r}, t) + C \tag{1.145}$$

As a result, the exact time-dependent density can be calculated as

$$\rho(\mathbf{r}, t) = \sum_i |\phi^{KS}(\rho(\mathbf{r}, t)|^2 \tag{1.146}$$

with aid of the fictitious Kohn-Sham orbitals, arising from a time-dependent version of the Kohn-Sham equations.

$$i\frac{\partial}{\partial t} |\phi_p^{KS}(\mathbf{r}, t)\rangle = \hat{f}^{KS}(\mathbf{r}, t) |\phi_p^{KS}(\mathbf{r}, t)\rangle \tag{1.147}$$

$$f^{KS}(\mathbf{r}, t) = h(\mathbf{r}) + v^{KS}(\mathbf{r}, t) \tag{1.148}$$

$$v^{KS}(\mathbf{r}, t) = v_s(\mathbf{r}, t) + v_{ext}(\mathbf{r}, t) \tag{1.149}$$

$$v_s(\mathbf{r}, t) = v_{eN}(\mathbf{r}, t) + v_H(\mathbf{r}, t) + v_{xc}(\mathbf{r}, t) \tag{1.150}$$

$$v_{xc}(\mathbf{r}, t) = \frac{\delta E_{xc}[\rho]}{\delta \rho(\mathbf{r}, t)} \tag{1.151}$$

Note how I have decomposed the time-dependent Kohn-Sham potential into the contributions inherent to the fictitious non-interacting system (v_s) which depend on time through the time-dependent density, and v_{ext} which is an external time dependent perturbation. A crucial difficulty up to this point is that the exchange-correlation potential in Eq. 1.151 depends in the density at all points in time. Yikes!

The adiabatic approximation

In a crucial step towards practical TDDFT calculations, the adiabatic approximation sets the exchange-correlation functional in Eq. 1.151 to that of the ground state, thus robbing the exchange-correlation kernel of any dependence on the frequency. As a result, this allows to develop a theory for excited states using ground state exchange-correlation functionals. How could such an approximate theory perform in predicting excitation energies? As will be elaborated shortly, very well for valence excited states but not so well for core excited states.

Linear response of the density and the Casida equations

When the external perturbation is small, response theory provides a useful construct to learn something about the system. The time and space non-local susceptibility of the full, interacting system

$$\chi(\mathbf{r}, \mathbf{r}', t - t') = \left.\frac{\delta\rho(\mathbf{r}, t)}{\delta v_{ext}(\mathbf{r}', t')}\right|_{\rho_0} \tag{1.152}$$

encodes the linear response of the density with respect to the external field. Defining the susceptibility of the non-interacting time-dependent Kohn-Sham density and the time-dependent exchange-correlation kernel

$$\chi_s(\mathbf{r}, \mathbf{r}', t - t') = \left.\frac{\delta\rho(\mathbf{r}, t)}{\delta v_s(\mathbf{r}', t')}\right|_{\rho_0} \tag{1.153}$$

$$f_{xc}(\mathbf{r}, \mathbf{r}', t - t') = \frac{\delta v_{xc}(\mathbf{r}, t)}{\delta\rho(\mathbf{r}', t')} \tag{1.154}$$

allows to write the interacting susceptibility in the following form (after Fourier-transform into the frequency domain). [111]

$$\chi(\mathbf{r}, \mathbf{r}', \omega) = \chi_s(\mathbf{r}, \mathbf{r}', \omega) + \int d\mathbf{r_1} d\mathbf{r_2} \chi_s(\mathbf{r}, \mathbf{r_1}, \omega) \times \left[\frac{1}{|\mathbf{r_1} - \mathbf{r_2}|} + f_{xc}(\mathbf{r_1}, \mathbf{r_2}, \omega)\right] \chi(\mathbf{r}, \mathbf{r}', \omega) \tag{1.155}$$

The poles of the susceptibility correspond to the excitation energies of the system. Writing this operator in matrix form in the basis of Kohn-Sham configurations results in the Casida

equations and allows us to extract the poles of the susceptibility as the solutions to a pseudo-eigenvalue problem. [112]

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{bmatrix} \begin{pmatrix} \mathbf{X}_k \\ \mathbf{Y}_k \end{pmatrix} = \omega_k \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{bmatrix} \begin{pmatrix} \mathbf{X}_k \\ \mathbf{Y}_k \end{pmatrix} \tag{1.156}$$

The elements of the A and B matrices given above come out to be

$$A_{ia,\,jb} = \delta_{ij}\delta_{ab}(\epsilon_a - \epsilon_i) + \langle ia| f_H + f_{\mathrm{xc}} |jb\rangle \tag{1.157}$$
$$B_{ia,\,jb} = \langle ia| f_H + f_{\mathrm{xc}} |jb\rangle \tag{1.158}$$

The Tamm-Dancoff approximation, which amounts to setting the B matrix to 0, turns out to provide results close to those from full TDDFT. [113]

Performance for valence and core excitations

So, how does TDDFT under the aforementioned approximations perform in predicting excitation energies? For the Quest data set, and with modern exchange correlation functionals, excellently. [106] The concept of Jacob's ladder, introduced for the performance in DFT in the ground state, is recovered for excitation energies so that more sophisticated functionals perform better. The best-performing meta-GGAs and hybrid-GGAs in the study of Liang *et al.* afford an RMSE of about 0.5 and 0.3 eV.

As with CIS, the story is a lot different for core excitations. Under the core-valence separation scheme brought up previously, the errors with even the best-performing functionals in Ref [106] are on the order of 10s of eV for K-edge core excitations on main group elements. [114] The errors climb to 100s of ev for K-edge excitations on third-row elements. TDDFT theory is exact prior to the introduction of the adiabatic and Tamm-Dancoff approximations and, in this case, errors for core excitations are likely due to the former. Two manifestations of the errors in core excitations due to the approximations invoked for practical TDDFT are 1) are the lack of orbital relaxation, in common with CIS and 2) the self-interaction error of an electron exciting the core with a residual image of itself. In the last two decades, specialized short-range corrected (SRC) functionals including a large amount of short-range exact exchange were designed to alleviate the latter and perform significantly better than standard functionals, to the detriment of other properties. [115–122] More recent approaches involving alternative references will be elaborated on further on.

1.3.4 Equation-of-motion coupled-cluster

General framework

The equation of motion (EOM) formalism, within the coupled-cluster framework, provides a means to construct excited states with electron correlation. [123, 124] Solving the CC equations for the ground state (Section 1.2.4) defines the similarity-transformed Hamiltonian

in Eq. 1.110. The right eigenvectors and eigenvalues of $\bar{H}$ within a subspace of excited configurations define the EOM-CC excited-state wave functions $|R_y\rangle$ and excitation energies ΔE_y.

$$(\bar{H} - E^{CC})\,|R_y\rangle = \Delta E_y\,|R_y\rangle \tag{1.159}$$

Naturally, the EOM-CC eigenvectors are linear combinations of the excited configurations within the EOM subspace (usually, but not necessarily, truncated in the same manner as for the ground state). For example, within the EOM-CCSD model, the yth right eigenvector takes the form

$$|R_y\rangle = \hat{R}_y\,|\varnothing\rangle \tag{1.160}$$

$$= \left(\sum_{ia}(r_y)_i^a\hat{\alpha}_i^a + \frac{1}{4}\sum_{ijab}(r_y)_{ij}^{ab}\hat{\alpha}_{ij}^{ab}\right)|\varnothing\rangle \tag{1.161}$$

While this is identical in construction to a CISD wave function the key here is, in contrast to the CI model, correlation is already baked into the effective Hamiltonian via the similarity-transformation. This difference is more evident when we write the EOM-CCSD right eigenvectors for the un-transformed Hamiltonian.

$$|R_y\rangle = \hat{R}_y\,|\phi^{\text{CCSD}}\rangle \tag{1.162}$$

$$= \hat{R}_y e^{\hat{T}}\,|\varnothing\rangle \tag{1.163}$$

An unattractive consequence of the EOM-CC model arises from the fact that the similarity-transformed Hamiltonian is not Hermitian, and thus there is a corresponding set of left-eigenvalues that are not related to the right-eigenvalues by complex conjugation

$$\langle L_x|\,(\bar{H} - E^{CC}) = \langle L_x|\,\Delta E_x \tag{1.164}$$

However, the left and right eigenvectors do form a bi-orthogonal set

$$\langle L_x|R_y\rangle = \delta_{xy} \tag{1.165}$$

Solving the separate equations for the left eigenvectors is required when properties of the excited states are desired, since these are expectation values of operators. When only excitation energies are sought, only the equations for the right eigenvectors are solved.

EOM-CC solvers and performance for valence and core excitations

In theory, the EOM formalism is similar to CI theory but the similarity-transformed Hamiltonian takes the place of the proper Hamiltonian. Therefore, the Davidson algorithm can be adapted to EOM-CC and is of crucial importance to make calculations on realistic systems tractable. [123] EOM-CC methods are generally well-regarded for their accuracy and robustness for electronic excitations. The most economical, EOM-CCSD, performs admirably for

valence excitations and there has been a considerable amount of work in designing accurate and efficient approximations to extend its approach to large systems. [125–129] Higher-level truncations such as EOM-CCSDT, EOM-CCSDQ, etc. are too computationally demanding and often not needed for routine use. [130, 131] They serve instead to provide benchmark numbers, such as those in the Quest dataset mentioned earlier. [99–105]

The story is slightly different for core excitations, since the EOM framework suffers from a couple of problems. As for CIS and TDDFT, the first one is finding a way to target the interior eigenvalues of the similarity-transformed Hamiltonian; to this end, a variety of approaches were developed. [132–136] A challenge that some techniques implemented in the last decade faced is the fact that core excited states are resonances embedded in an Auger continuum. [132, 133] The earlier idea of the core-valence separation emerged as a successful solution to explicitly decouple the continuum from the core excited states, and therefore has become the preferred protocol to target core excitations. [107, 135–137]

It is worth noting, however, that the details of the CVS implementation may lead to differences on the order of eVs. The CVS scheme of Coriani *et al.* includes all excitations for the ground state and decouples the core excited states via projection from the valence states in the EOM component of the procedure. [135] With this model, EOM-IP-CCSD is in error by 1 - 3 eV for core ionizations of small organic molecules, and full triples are required to bring errors below 0.3 eV. [138] This is in stark contrasts to valence excitations of closed-shell single-reference molecules, where EOM-CCSD already achieves sub-eV accuracy per its performance on the Quest dataset. The root of this discrepancy is the same that affects CIS and TDDF: the strong orbital relaxation effects present in core excited states place a stronger-than-usual burden on the EOM-CC framework to describe correlation.

The frozen-core (fc)-CVS scheme of Vidal *et al.* employs the frozen-core approximation on the ground state CC wave function, and applies the EOM excitation operator exclusively on the core orbitals of interest. [136] As a result, the ground state wave function is under-correlated and, due to the resulting cancellation of errors, the fc-CVS EOM-CCSD approach performs better: typically, excitation energies are in sub-eV agreement with experiment. It would be interesting to see the performance of fc-CVS EOM-CCSDT; based on the results between the Coriani-style CVS-EOM-CCSD and -CCSDT for core ionizations, I expect the fc-CVS EOM-CCSDT results to deliver energies lower than experiment and errors equal to or larger than those of the -CCSD results.

In a final note, I mention the recent implementation of a spin-adapted version of EOM-CCSD to treat core excited states for triplet species. [139]

1.3.5   Orbital-optimization state-specific approaches

Response theories ultimately involve diagonalization of an effective Hamiltonian and results in multiple excited states within one calculation. The so-called state-specific methods provide an alternative approach that targets individual excited states in some manner. In general, state-specific approaches benefit from improved accuracy in producing excitation energies compared to their sibling response-based approach. As a result, state-specific ap-

proaches are appropriate for providing theoretical benchmark numbers, as well as for use in the interpretation of experimental spectra where high accuracy is required.
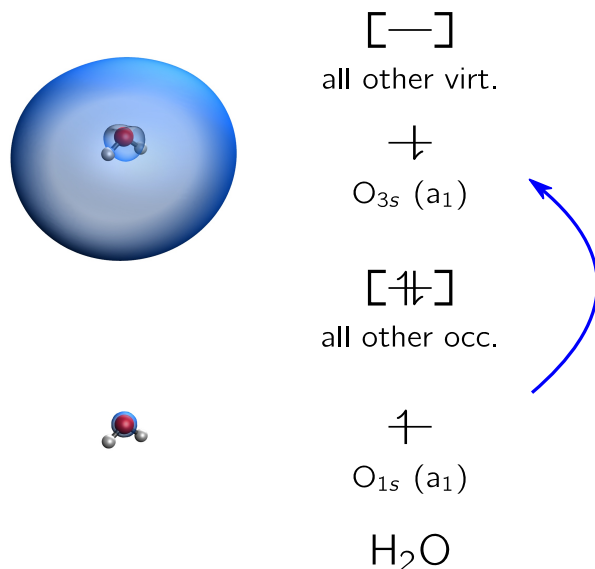
Figure 1.6: Non-Aufbau configuration of water for the $O_{1s} \rightarrow O_{3s}$ core excited state.

In payment for accuracy, the use of state-specific approaches forsakes convenience. State-specific methods become cumbersome for situations where numerous excited states are required such as calculations on large systems, when attempting to faithfully reproduce the state-dense Rydberg regime, or when employing molecular dynamics simulations to generate a large number of nuclear configurations to employ in spectral calculations.

Imagine, for example, the configuration corresponding to a $O_{1s} \rightarrow O_{3s}$ excitation in water (Figure 1.6) constructed from the canonical orbitals of the ground state. It turns out that carrying out the SCF procedure with this excitation as a starting point is able to locate a solution to the SCF equations that indeed resembles the core excited configuration. The energy difference in between the orbitals optimized for the ground state and the orbitals optimized for the ionized state defines the ΔSCF energy.

$$\Delta SCF(E) = E_0 - E(O_{1s} \rightarrow O_{3s}) \tag{1.166}$$

The physical appeal of orbital-optimized procedures lies in that the orbitals are re-shaped by the optimization procedure to describe the electronic nuances of the excited state itself as opposed to employing the orbitals tailored to the ground state, as response theories do. In other words, excited-state SCF optimization provides more appropriate orbitals to describe excited state densities. As it turns out, this is absolutely crucial for core-excited resonances because their density differs greatly from that of the ground state. The excitation energies delivered by orbital-optimized approaches is vastly superior to, say, traditional TDDFT.

Specialized SCF solvers

Several shortcomings must be addressed when employing orbital-optimized approaches. The first is that excited (non-Aufbau) solution to the SCF equations is stationary to orbital rotations by definition (i.e. has a zero orbital gradient) yet it is unstable with respect to the hole-filling orbital rotation (i.e. has an orbital Hessian with at least one negative eigenvalue). In other words, it is a (first- or higher-order) saddle point in orbital space. As a result it is difficult, if not impossible, to carry out the SCF procedure and converge to the correct state with standard SCF solvers like DIIS and GDM; variational collapse (landing in the Aufbau solution to the SCF equations) is the result more often than not. This problem is well-established and several procedures have been proposed to address it.

The maximum overlap method (MOM) establishes the orbitals occupied at each step in the SCF procedure to be the ones with maximum overlap to those of the previous iteration. [140] The MOM thus guides the SCF procedure to the desired solution, provided an initial guess sufficiently close to the desired SCF solution is available. Often, this may just be the target non-Aufbau configuration generated from the orbitals optimized to the ground state. A subsequent off-shoot of this method, the initial MOM (iMOM), uses the initial guess as the fixed template for the orbital occupation during each SCF iteration. [141] When targeting core excitations well-isolated energetically from other states, the MOM-based approaches are often sufficient for optimization. However, the Rydberg regime is dense and converging to the correct configuration may become difficult for MOM because of fortuitous overlaps between the current iterate and other states.

The square gradient minimization (SGM) procedure presented a more robust, though somewhat more costly, means to converging core-excited solutions. [142] SGM relies on the fact that, while non-Aufbau solutions feature instabilities in the orbital Hessian of the target Lagrangian $\mathcal{L}$ (i.e. the SCF energy under the appropriate constraints, Eq. 1.32) with respect to core-filling orbital rotations, the square of the orbital gradient

$$|\nabla_{\tilde{\theta}}\mathcal{L}|^2 \tag{1.167}$$

is positive semi-definite even at this point in orbital space. The problem of converging to a saddle point of the SCF energy in orbital space can be transformed into a minimization of the objective function Eq. 1.167 and is thus more robust. SGM, however, comes with the caveat of numerical instabilities arising from the ill-conditioned nature of the problem of optimizing the square of a gradient. Specifically, the optimization becomes in general more difficult as the condition number of the relevant hessian increases, and SGM *squares* the condition number of the orbital hessian.

One of the most recent SCF solvers for excited states, the state-targeted energy projection (STEP) method, imposes an energy level-shift on the virtual orbital from the perspective of a non-Aufbau configuration $\mathbf{Q}$. [143] Importantly, this virtual space includes the would-be occupied orbital for an Aufbau occupation.

$$\mathbf{F}^{'} = \mathbf{F} + \eta\mathbf{SQS} \tag{1.168}$$

Effectively, this level shifts transforms the desired non-Aufbau configuration into the lowest-energy configuration by raising the energy of the nominally-occupied orbital beyond that of the nominally unoccupied orbital. There are two limitations to STEP. The first is that the level-shift parameter η is arbitrary. In general, one needs to apply a level-shift strong enough to make the target hole the effective highest-occupied MO (HOMO) and then a bit more to securely avoid variational collapse. However, in the process, the occupied-virtual rotations necessary for convergence are supressed due to an increasing energy gap. The second limitation is robustness; while STEP often provides an improvement upon MOM, some scenarios STEP cannot handle. For example, I've found both STEP and MOM to be unable to converge 3p core holes in certain iron-containing compounds (Chapter 4). Only SGM manages to painfully converge them.

Spin contamination

The second problem is that excited states and ionized states are, more often than not, radical species. As elaborated on Subsections 1.2.2 and 1.2.2, open-shell species require CSFs in general so single-reference models feature qualitative failures. Only in the case when the desired species has all electrons "pointing" in the same direction (i.e. the maximum or minimum value of S_z for the target S^2 value) is a single determinant appropriate. For example, the singlet $O_{1s} \rightarrow O_{3s}$ excited state of water introduced earlier on, requires two-configurations by virtue of being a two-electron open-shell system (Eq. 1.70) As a result, the excited state energy afforded by ΔSCF, call it $^1E_{\text{USCF}}$, is said to be spin-contaminated because the sole configuration employed for the evaluation of the excited state energy can be obtained from a linear combination in between the singlet and triplet CSFs.

Relying on this same fact, Yamaguchi and coworkers proposed the approximate spin-projection scheme to estimate the energy of the proper singlet ($^1E_{\text{proper}}$) from a spin con-taminated configuration, provided the "degree of spin contamination" (i.e. $^1\langle S^2 \rangle_{\text{USCF}}$) and the energy of the spin-contaminant ($^3E_{\text{USCF}}$) are available. [144]

$$^1E_{\text{proper}} =^1 E_{\text{USCF}} + \alpha(^1E_{\text{USCF}} -^3 E_{\text{USCF}}) \tag{1.169}$$

$$\alpha = \frac{^1\langle S^2 \rangle_{\text{USCF}}}{^3\langle S^2 \rangle_{\text{USCF}} -^1 \langle S^2 \rangle_{\text{USCF}}} \tag{1.170}$$

The beauty of this approach lies in that, within the triply-degenerate two-electron open-shell triplet manifold, the $M_S = \pm 1$ states have all electrons pointing in the same direction so their energies can be reliably captured by standard USCF or ROSCF and be used as an exact replacement for the $M_S = 0$ triplet spin contaminant in ΔSCF. Note, however, that higher-spin contaminants (quintet, septet, ...) may also be present and the AP, as formulated above, does not address the contamination due to those. In practice, nonetheless, core excited states of sinlge-excitation character are often very-well treated with the AP.

The restricted open-shell Kohn-Sham (ROKS) theory goes a step further by optimizing the perfectly spin-purified energy provided by applying the spin-projection procedure on

the spin-contaminated singlet energy with the triplet energy evaluated with the same set of restricted open-shell orbitals. [145] To put it another way, ROKS minimizes the energy of a single singlet CSF. ROKS achieves an accuracy on the order of 0.3 - 0.4 eV for the K-edge of main group elements and, with a perturbative treatment of spin-orbit coupling (SOC), L-edges of second-group elements. [146] In a subsequent study, it was found that accounting for scalar relativistic effects via the spin-free exact two-component (X2C) model [147] extended the applicability of ROKS to the K-edge of third group elements and the first few transition metals before higher-order relativistic effects take hold. [148]

The AP philosophy can be extended to general spin-recoupling schemes to provide spin-pure energies for arbitrary open-shell systems. Hait *et al.* made use of this approach to evaluate the three-electron open-shell (3eOS) and 4eOS core excited states for doublet, triplet, and open-shell singlet radicals. [149, 150] However, to the best of my knowledge, the exploration of ROKS (as opposed to AP-ΔSCF) has remained confined to 2eOS.

Several other state-specific frameworks have arrived at similar solutions to the multi-configurational character of low-spin open-shell excited states. Originally designed to describe the non-adiabatic coupling between several electronic states, the multi-state density functional theory (MS-DFT) was recently adapted to predict the doublet 3eOS core excited states of radicals with the attractive feature of addressing the correlation-double counting inherent to the use of DFT orbitals in CI-theory. [151–155]

Excited-state mean-field (ESMF) theory is a more general state-specific framework that invokes an excited-state variational principle to minimize the energy of an excited state. [156–159]

$$W = \frac{\langle \Phi | (\omega - H)^2 | \Phi \rangle}{\langle \Phi | \Phi \rangle} \tag{1.171}$$

In practice, an approximation to Eq. 1.171 is used to avoid the difficulties associated with squaring the Hamiltonian. For the purposes of describing singly-excited statates, an anzats of the form below is appropriate. [160]

$$|\Phi^{ESMF}\rangle = e^X \left(c_0 \, |\varnothing\rangle + c_i^a \, |\Phi_i^a\rangle \right) \tag{1.172}$$

X is an orbital rotation operator of the form in Eq. 1.61 and it is the key to provide orbitals that properly describe the excited state. ESMF theory is, rigorously, more appealing than ROKS for two reasons. First, ESMF includes the ground state configuration in the anzats which alleviates the strong overlaps between the ground and excited wave functions that often plagues ROKS. [161] Second, the ESMF anzats in Eq. 1.172 allows to describe singly excited states which feature genuine configuration interaction beyond that required by spin-symmetry (i.e., those that require more than one NTO CSF pair). This technique was employed by Garner and Neuscamman to predict the core excitation energies of a series of small molecules with errors comparable to that of ROKS. [162] Crucially, ESMF is a mean-field method void of electron correlation so its accuracy, on the order of RMSE = 0.6 eV, is remarkable and speaks to the dominant role of orbital relaxation for the description of core excited states.

The inconvenience of state-specific methods

Having addressed the difficulties of solving for core-excited configurations and spin contamination, the main disadvantage that remains for state-specific methods in general is more technical in nature: they are inconvenient to use. In the case of AP-ΔSCF or MS-DFT, a number of unrestricted configurations (two, three, and six for 2eOS singlets, 3eOS doublets, and 4eOS singlets, respectively) must be optimized independently to calculate spin-pure excited states. Asserting that the configurations indeed correspond to the same set of spatial orbitals presents a challenge to automation procedures. While progress in this area has been accomplished for 2eOS core excited states via AP-ΔSCF, [163] no such automation procedure has been designed for 3eOS and 4eOS excited states.

Even in the case when spin-adaptation is incorporated by construction in the method, such as in ROKS or ESMF, difficulties in convergence and the fact that each state must still be optimized individually makes state-specific approaches cumbersome for systems of moderate size and for generating a large number of spectra for different nuclear configurations. Therefore, full-spectrum "state-universal" methods are preferable relative to state-specific approaches when their accuracy is sufficient for the target purpose.

1.3.6   Correlated state-specific approaches

What happens if you employ an state-specific, orbital-relaxed reference as the starting point for a correlated calculation? This is the essence of the ΔMP, ΔCC, and ΔCI approaches, as well as the correlated extensions to ESMF theory, which have gathered interest in the last two decades and especially in the last few years.

As it turns out, excited SCF solutions are often a better reference than the ground state for finding alternative solutions to the CC equations, which in turn are reasonable approximations to the true excited states. [164] The appeal of these approaches for core excited states lies in that explicit SCF re-optimization takes care of the strong orbital relaxation and allows correlated methods such as MP-PT and CC to focus on addressing the remaining dynamic correlation of a system.

The problems that hinder orbital-optimized state-specific methods, namely the possibility of variational collapse and spin contamination, carry over to their correlated counterparts relying on single-reference formalisms. Since these methods are central to this thesis, the approaches taken to address variational collapse and spin contamination have a special place in the Introduction of Chapter 2 (Section 2.1) and will be elaborated on there.

1.3.7   Orbital-optimized, full-spectrum approaches

Is there a way to reconcile the convenience of diagonalization-based approaches with the explicit orbital relaxation from which state-specific methods benefit from?

Static exchange and non-orthogonal configuration interaction singles

The first of attempts at doing this build upon the work of Hunt and Goddard III in 1969, where they realize that the singly-excited configurations out of an optimized n-electron configurations more appropriately describe the electron-attached states for a $(n + 1)$-electron system. [165] Leveraging this idea, the authors use the optimized $(n - 1)$-electron configurations to provide better approximations to the excited states of the n-electron system in what has historically been called the static exchange (STEX) approach. In the modern nomenclature of Fock space approaches to ionized and electron attached states, it is better thought of as electron affinity CIS (EA-CIS), as the CIS is performed after adding back one electron to the ionized orbital reference determinant. Ågren and co-workers adapted EA-CIS to treat core excited states by using the orbitals optimized for the core-ion, thus conferring the reference orbitals for the excited states with the much needed core-hole relaxation, yet allowing the whole set of core excited states to be obtained via CI theory. [166–168] Ever since, EA-CIS has been used to calculate full-spectrum NEXAS with success for decades.

Note that EA-CIS for core excitations, like the orbital-optimized approaches, requires the localization of the core orbitals prior to re-optimization when they are delocalized due to the symmetry of the molecule. [169] For example, the canonical nitrogen core orbitals of the N_2 molecule, out of respect to the $D_{\infty h}$ symmetry of the molecule, are delocalized over the molecule in σ and σ^* fashion. In the work of Katherine Oosterbaan (one of my grad. school mentors) and coworkers, the EA-CIS approach was augmented to including the coupling between excited states out of diferent core orbitals in the non-orthogonal configuration interaction singles (NOCIS) approach. [109] Since each core-ionized reference is optimized separately, the set of orbitals for each are non-orthogonal to the other and the CI matrix elements must be constructed with non-orthogonal configuration techniques. All in all, EA-CIS and / or NOCIS achieves an RMSE of about 1.5 eV on K-edge excitations for a comprehensive sample of single-reference organic molecules. [109, 114] While not at the sub-eV accuracy afforded by state-specific methods, NOCIS provides an order of magnitude in improvement over the analogous CIS with ground state orbitals.

EA-CIS and NOCIS were subsequently generalized to treat one-electron open-shell radical species, such as CO^+ or $N_2{}^+$. [170, 171] In a further step, Oosterbaan and co-workers designed a NOCIS strategy for high-spin triplet two-electron open-shell references and introduced the one-center (1C) approximation to NOCIS, where the excitations arising from different core orbitals are decoupled. [172] 1C-NOCIS and EA-CIS are thus identical for closed-shell molecules but they differ for radicals, for which NOCIS employs different sets of orbitals for different class of excited states. Specifically, NOCIS explicitly optimizes the electron-repairing core excitations via ΔSCF, while using core ionized orbitals for the remaining excited states.

A project I was a part of but which will not be included in this thesis involved generalizing the EA-CIS approach to DFT. The resulting method, called electron-affinity (EA)-TDDFT, achieved an 0.5 RMSE across a large data-set of 132 core excitation energies on organic molecules of varied character for first- and second-row elements (σ^*, π^*, Rydberg). [114] A

similar idea was proposed by Hait *et al.*, where the ROKS energies with a specific functional are evaluated with core-ionized references in what they called the ROKS(STEX) approach. [173]

Spin-flip methods

In general, spin flip (SF) methods in quantum chemistry employ a high-spin open-shell reference, instead of the ground state, as a starting point to access excited states via spin-flip operators. [174, 175] Park and co-workers employed a multi-reference (MR)-SF-TDDFT approach, where both $M_s = 1$ and $M_s = $ -1 core-hole triplet references are employed to target both the closed-shell ground state as well as its core excited states. [176] Notably, this approach also allows a description singlet excited states of valence character and partially describe the core excited states atop them, which result in configurations with four unpaired electrons. Designing a theory for these exotic states is the main concern of Chapter 3 in this thesis. Such excited states are relevant to modern UV-vis pump / X-ray probe spectroscopies because the final state in such experiments is a core excitation on top of a valence excitation in exactly that way.

Transition-potential methods

The transition-potential (TP)-SCF method seeks to approximate the ΔSCF energy of a specific excited state as an orbital energy difference for a configuration optimized to a fractional number of electrons in the open shells of interest (a core orbital and a virtual orbital). TP-SCF has been used for decades for core excitations and continues to be studied today. [177–180]

A nice fact about TP-SCF is that it provides a reference striking a balanced description of the ground state and the core excited states. Simons and Matthews leveraged this fact to employ TP-SCF references for use in the EOM-CC framework. [181] Naturally, the TP-CC method inherits some of the advantages of both state-specific methods - orbital relaxation - while retaining the advantages of EOM-CC: inherent spin-adaptation of the excited states, a full spectrum with a single calculation, and straightforward transition properties. The cost to pay comes from relying on a deteriorated description of the ground state relative to standard CC, controlled by tuning the fractional occupation number of the core orbital in the underlying reference. Even though this renders the model arbitrary, to some extent, Simons and Matthews have carried out a study to find an optimal core occupancy parameter transferable across edges of the same element, making this a promising method for reliable and affordable high-accuracy wave function X-ray calculations. [182]

1.4 Outline

Below are the abstracts of the following chapters.

Chapter 2

The content and figures of this chapter are reprinted or adapted with permission from **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. [183]

We investigate the use of orbital-optimized references in conjunction with CCSD for the study of core excitations and ionizations of 18 small organic molecules, without the use of response theory or equation-of-motion (EOM) formalisms. Three schemes are employed to successfully address the convergence difficulties associated with the coupled-cluster equations, and the spin contamination resulting from the use of a spin symmetry-broken reference, in the case of excitations. In order to gauge the inherent potential of the methods studied, an effort is made to provide reasonable basis set limit estimates for the transition energies. Overall, we find that the two best-performing schemes studied here for ΔCCSD are capable of predicting excitation and ionization energies with errors comparable to experimental accuracy. The proposed ΔCCSD schemes reduces statistical errors against experimental excitation energies by more than a factor of two when compared to the frozen-core core-valence separated (fc-CVS) EOM-CCSD approach - a successful variant of EOM-CCSD tailored towards core excitations.

Chapter 3

The content and figures of this chapter are reprinted or adapted with permission from **Arias-Martinez, J. E.**; Wu, H.; Head-Gordon, M. "Generalization of one-center non orthogonal configuration interaction singles to open shell singlet reference states: Theory and application to valence-core pump-probe states in acetylacetone" *arXiv e-prints*, **2023**, arXiv:2310.09684. [184]

We formulate a one-center non-orthogonal configuration interaction singles (1C-NOCIS) theory for the computation of core excited states of an initial singlet state with two unpaired electrons. This model, which we refer to as 1C-NOCIS two-electron open-shell (2eOS), is appropriate for computing the K-edge near-edge X-ray absorption spectra (NEXAS) of the valence excited states of closed-shell molecules relevant to pump-probe time-resolved (TR) NEXAS experiments. With inclusion of core hole relaxation effects and explicit spin adaptation, 1C-NOCIS 2eOS requires mild shifts to match experiment, is free of artifacts due to spin contamination, and can capture the high-energy region of the spectrum beyond the transitions into the singly occupied molecular orbitals (SOMO). Calculations on water and thymine illustrate the different key features of excited-state NEXAS, namely the core-to-SOMO transitions as well as shifts and spin-splittings in the transitions analogous to those

of the ground state. Finally, simulations of the TR-NEXAS of acetylacetone after excitation onto its $\pi \rightarrow \pi^*$ singlet excited state at the carbon K-edge - an experiment carried out recently - showcases the ability of 1C-NOCIS 2eOS to efficiently simulate NEXAS based on non-adiabatic molecular dynamics simulations.

Chapter 4

The content and figures of this chapter are reprinted or adapted with permission from Troß, J.; **Arias-Martinez, J. E.**; Carter-Fenk, K.; Cole-Filipiak, N. C.; Schrader, P.; McCaslin, L. M.; Head-Gordon, M.; Ramasesha, K. "Femtosecond core-level spectroscopy reveals signatures of transient excited states in the photodissociation of $Fe(CO)_5$" *ChemRxiv e-prints*, **2023**, 10.26434/chemrxiv-2023-m1ggn. [185]

Excitation of iron pentacarbonyl (IP), a prototypical photocatalyst, at 266 nm causes sequential loss of two CO ligands in the gas phase, creating catalytically active, unsaturated iron carbonyls. Thus far, the electronic states involved in the dissociation have eluded experimental observation, hindering a comprehensive understanding of IP photochemistry. Using femtosecond extreme ultraviolet transient absorption spectroscopy near the Fe $M_{2,3}$-edge, we present the first spectroscopic characterization of valence electronic dynamics during IP photodissociation. Informed by electronic structure calculations, we uncover the spectroscopic signatures of the intertwined structural and electronic evolution among the manifold of metal-centered excited states during first CO loss from IP on a 100-fs timescale. Furthermore, spectroscopic signals associated with the formation of $Fe(CO)_4$ on its lowest singlet excited state and in structures fluctuating between C_{2v} and C_{3v} geometries, and its subsequent picosecond dissociation to $Fe(CO)_3$ in the C_s geometry, are corroborated using quantum chemistry calculations.

Projects I was a part of that are not included in this thesis

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. Chem. Phys. Lett.*, **2022**, 13, 9664-9672. [114]

The lack of particle–hole attraction and orbital relaxation within time-dependent density functional theory (TDDFT) lead 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 the 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 the 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.

Talbot, J. E.; **Arias-Martinez, J. E.**; Cotton, S. J.; Head-Gordon, M. "Fantastical excited state optimized structures and where to find them" *J. Chem. Phys.*, **2023**, 159, 171102. [186]

The quantum chemistry community has developed analytic forces for approximate electronic excited states to enable walking on excited state potential energy surfaces (PES). One can thereby computationally characterize excited state minima and saddle points. Always implicit in using this machinery is the fact that an excited state PES only exists within the realm of the Born-Oppenheimer approximation, where the nuclear and electronic degrees of freedom separate. This work demonstrates through ab initio calculations and simple nonadiabatic dynamics that some excited state minimum structures are fantastical: they appear to exist as stable configurations only as a consequence of the PES construct, rather than being physically observable. Each fantastical structure exhibits an unphysically high predicted harmonic frequency and associated force constant. This fact can serve as a valuable diagnostic of when an optimized excited state structure is non-observable. The origin of this phenomenon can be attributed to the coupling between different electronic states. As PESs approach one another, the upper surface can form a minimum that is very close to a near-touching point. The force constant, evaluated at this minimum, relates to the strength of the electronic coupling rather than to any characteristic excited state vibration. Nonadiabatic dynamics results using a Landau-Zener model illustrate that fantastical excited state structures have extremely short lifetimes on the order of a few femtoseconds. Their appearance in a calculation signals the presence of a nearby conical intersection through which the system will rapidly cross to a lower surface.

Chapter 2

Accurate core excitation and ionization energies from a state-specific coupled-cluster approach

2.1   Introduction

While excited state electronic structure has found tremendous success for describing valence excitations of molecular systems, the same cannot be said for core excitations due to two main challenges. The first is the fact that core excited states lie high beyond the valence continuum. Developments such as the core-valence separation (CVS) scheme have allowed response theories to access core excited states and thus address this difficulty (Section 1.3.2). The second challenge, related to the strong orbital relaxation caused by a removal from an electron in the core orbital, is more difficult for response theories to address. For example, traditional TDDFT requires specialized functionals to avoid incurring errors on the order of tens of eVs for the K-edge of the main group elements (Section 1.3.3). Even EOM-CC requires truncation at full triples to achieve performance on par to EOM-CCSD for the valence regime (Section 1.3.4). Circumventing the shortcomings of response theories is an active area of research with exciting recent advances (Section 1.3.7).

Inspired by the spectacular performance of state-specific orbital-optimized methods for core transitions (Section 1.3.5), this project sought to explore how well the correlated state-specific methods could compare. [146, 148, 162] ΔMP2, ΔCCSD, ΔCCSD(T), and even ΔCCSDT have been used in the past decade to describe core ionized states. [178, 187–194] However, the ESMP2 work of Garner and Neuscamman and the TD-CCSD results of Matthews had been the only to explore core excited states. [162, 195] Aside from the variational collapse and divergence problem, the relative scarcity of state-specific studies on core excitations arises due to the multi-configurational nature of the excited states, which prevents the immediate application of single-reference theories without modifications.

Owing to the simple nature of the multi-configurational character of singly core excited

states of closed shell systems (namely, a two-determinant CSF) the objective of this paper
is to assess the use of single-reference ΔCC for the prediction transition energies in the core
regime. We present three approaches to address the ill-behaved amplitudes and spin con-
tamination inherent in single-reference CCSD for state-specific optimization of core excited
states employing spin-symmetry-broken references. Two of them employ Yamaguchi's AP
approach, [144] while the third one instead enforces correct spin symmetry at the ROHF
level by constraining the amplitude of the double substitution that flips the spins of the two
half-occupied orbitals to $+1$ for singlet and -1 for triplet states.

A comparison of these approaches is made against successful core excited state theories,
ROKS(SCF) and fc-CVS-EOM-CC, [136, 146] with the ultimate judge being the experimen-
tal results. As observed in this study, the best ΔCCSD models significantly outperforms fc-
CVS-EOM-CCSD while retaining its $\mathcal{O}(N^6)$ scaling, with N being the size of the one-electron
basis set employed. Furthermore, unlike fc-CVS-EOM-CC, it does not rely on cancellation
of errors. The protocols presented here are well-defined in that only the molecule and the
transition of interest needs to be specified - the proper ground state CC wave function and
energies are used as is and no compromise in the excited state wave function is made either.
The energetic differences between the singlet and triplet core excited states, presumed to be
accurate enough to make a statement about them, are presented.

An effort is made to reach basis set convergence for all methods in order to exclude
this factor from the discussion as much as possible and focus on their inherent performance.
Despite the computational demands of approaching the basis set limit (BSL) for CC methods
constraining us to molecules with at most two heavy atoms, the data set is diverse in terms
of the elements targeted (Be, C, N, O, F, Ne) and in terms of the excited state character
(σ^*, π^*, Rydberg). In total, a set of 21 excitations and 18 ionizations on 18 small closed-shell
organic molecules is used.

Our focus is on reporting excitation energies obtained through different proposed schemes
within the ΔCC framework, which we believe could be useful for providing theoretical bench-
mark numbers for core excited and ionized states. At present, our work does not extend the
discussion of ΔCC to compute transition properties. Obtaining such properties would be
cumbersome and expensive due to, in part, the use of different sets of amplitudes for the bra
and ket CC states. As pointed out in Ref. [192], a potentially useful strategy to circumvent
this exponential cost would be to use linearized wave functions obtained from the CC am-
plitudes from either the ground or core excited states, but we did not explore this further in
our study.

2.2 Background

Variational collapse and divergence issues

For a set of orbitals, not necessarily canonical, the CCSD amplitude equations take the form below. [196]

$$D_i^a t_i^a = F_{ia} + w_i^a(T_1, T_2) \tag{2.1}$$

$$D_{ij}^{ab} t_{ij}^{ab} = \langle ij | |ab \rangle + w_{ij}^{ab}(T_1, T_2) \tag{2.2}$$

The terms $w_i^a(T_1, T_2)$ and $w_{ij}^{ab}(T_1, T_2)$ in Eqs. 2.1 and 2.2 contain terms that are linear and higher in T_1 and T_2 separate from the orbital energy differences, D_i^a and D_{ij}^{ab} defined below.

$$D_i^a = \varepsilon_i - \varepsilon_a \tag{2.3}$$

$$D_{ij}^{ab} = \varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b \tag{2.4}$$

ε_p are the orbital energies themselves (Eq. 1.33). D_i^a and D_{ij}^{ab} will always be negative when employing a ground state reference and, in the absence of strong correlation, they are large enough to make the T amplitudes well behaved and much smaller in magnitude than 1. On the other hand, state-specific CC correlates a non-Aufbau SCF reference where the presence of a virtual orbital with a large negative energy - the core hole - allows for denominators D_i^a and D_{ij}^{ab} to be positive. For example, imagine an ionized SCF reference with a core hole in orbital $\bar{h}$ (we reserve the indexes h and $\bar{h}$ for the alpha core orbital and the beta core orbital). In the case of single excitations, the denominator $D_i^{\bar{h}}$ becomes positive when the occupied orbital has a higher orbital energy than the core virtual

$$\varepsilon_i > \varepsilon_{\bar{h}} \tag{2.5}$$

The condition in Eq. 2.5 holds unless there are other core orbitals of lower orbital energy. In the case of double excitations, $D_{ij}^{\bar{h}b}$ will be positive when

$$\varepsilon_i + \varepsilon_j - \varepsilon_b > \varepsilon_{\bar{h}} \tag{2.6}$$

One scenario where this happens is when the excitation $a_{\bar{h}}^\dagger a_i$ involves a valence occupied orbital and the excitation $a_b^\dagger a_j$ involves only valence orbitals. Positive denominators represent the desire of the standard CC formalism to seek the core-filling configuration and it is the manifestation of variational collapse within the context of CC theory. Variational collapse aside, the presence of a core hole in the reference can lead to numerical difficulties in solving for the T amplitudes and the divergence of perturbation theory when the orbital energies conspire to make $\varepsilon_i + \varepsilon_j - \varepsilon_b \approx \varepsilon_{\bar{h}}$, rendering $D_{ij}^{ab} \approx 0$. This problem is related with the continuum-coupling phenomena observed in EOM-CC that inspired the different CVS schemes. [138]

To address the divergence problem arising from positive and near-zero denominators in the ΔCC calculations of core ionizations Zheng *et al.* proposed to exclude the virtual core

orbital from the correlation treatment. [191, 193] Lee *et al.* adopted a similar strategy that excludes the doubly-vacant core orbital all together when studying double core-hole excitations. [192] Importantly, Zheng *et al.* found the missing correlation to be relevant for accurate core ionizations and uses estimates from fully-correlated CC calculations with decreasing denominator thresholds to account for it.

Spin contamination

In general, open-shell species require CSFs as opposed to single determinants for a spin-pure description (Sec. 1.2.2). As mentioned before, a single determinant is only appropriate for configurations where all electrons "pointing" in the same direction. Core ionized states of closed-shell systems are thus perfect cases to be treated by these models and they have been studied via ΔMP2 [178, 187–190] and, more recently, ΔCC.[191–193]

Singlet core excited-states of closed-shell systems require a two-determinant CSF and have therefore been more scarcely studied with correlated methods relying on a non-Aubfau reference. [187, 195] In a notable study, closely-related to this work, Matthews applied a direct two-determinant (TD) CCSD protocol to study singlet core excited-states with a scheme to address the "dangerous denominators" elaborated on previously. [195, 197–199] TD-CCSD performs well, with a MAE of 0.10 eV and RMSE of 0.11 eV against the Coriani-style CVS-EOM-EE-CCSDT for the three lowest lying core excitations of HCN, CO, NH_3, and H_2O.

2.3 Computational details

A development version of Q-Chem 5.4 was used for all calculations. Experimental geometries available on the NIST computational database were used throughout this work. [200] An atomic relativistic correction calculated via the Douglas-Kroll-Hell method, found to be nearly independent of basis-set and molecule for the main group elements, is added to all calculations (0.012, 0.09, 0.18, 0.34, 0.57, and 0.91 eV for Be, C, N, O, F, and Ne). [201]

We make use of three different kinds of (beta) core excited reference orbitals: (i) open-shell, symmetry-broken $M_S = 0$ references for the calculation of the singlet core excited states; open-shell, (ii) spin-pure triplet $M_S = 1$ references for the AP approach, when needed; and (iii) open-shell, spin-pure $M_S = \frac{1}{2}$ doublet references for the calculation of core ionized states. In the case of the spin-pure triplet and pure doublet references, standard ROHF is used in conjunction with the MOM algorithm (See "Specialized SCF solvers" in Section 1.3.5). The use of unrestricted orbitals for the symmetry-broken reference was found to be detrimental to some of our ΔCC schemes so ROKS(HF) orbitals, followed by a Fock-build for the broken-symmetry singlet state and subsequent pseudocanonicalization, were employed instead.

For two of the three schemes of ΔCC we employ, the calculated singlet excited states are spin contaminated; the AP method is used to estimate the spin-pure excitation energies.

Aside from the amplitudes excluded in the different schemes, the CC calculations of both the ground and excited states are all-electron.

Our best attempt was made at comparing the excitation or ionization energies near their BSL values. To that end, different procedures involving specialized basis sets were employed for obtaining an approximate BSL for the different methods. The aug-pcX-3 (heavy)/ aug-pcseg-2 (hydrogen) basis was used to approximate the BSL for the ROKS(SCF) calculations. [202] A (99, 590) Euler-Maclaurin-Lebedev grid was used for the computation of the exchange-correlation integrals for the ROKS(SCAN) calculations. The aug-ccX-nZ (heavy) / aug-cc-pVTZ (hydrogen) bases, [203] extrapolated using the two-point X^{-3} scheme with n = T, Q, were used to approximate the BSL for the EOM-CC calculations. [204, 205] As noted in a recent study, such an extrapolation scheme is appropriate for core excitations via EOM-CC. [206] All ROKS(SCF) and EOM-CC calculations were also run with the standard Dunning aug-cc-pCVXZ (X = D, T, Q) family of bases and a slower convergence towards a similar BSL value was observed. [207, 208]

Of the basis sets available, none were designed with both explicit orbital relaxation via SCF and correlation with wave function methods in mind. We used the TQ-extrapolated aug-cc-pCVXZ (heavy) / aug-cc-pVDZ (hydrogen) numbers as the best BSL estimate of the correlated Δ calculations.

The only exception to these choices of basis set was for the calculated Rydberg excitations in Ne. As expected for a full-fledged Rydberg excitation, significant differences between the aug-cc-pCVXZ and its doubly-augmented counterparts were observed in this case. The BSL core excited states for this atom are given by the d-aug-cc-pCV5Z for ROKS(SCF), Q5-extrapolated d-aug-cc-pCVXZ for EOM-CC, and TQ-extrapolated d-aug-cc-pCVXZ for the correlated Δ methods. No severe difference of a similar sort was found in any other molecule studied in this data set, including the rest of the isoelectronic ten electron series (see SI of the paper).

2.4   Approaches to inclusion of core-valence correlation

To motivate the need for the schemes presented in the following subsections, we begin by exploring the behavior of the Δ-correlated methods with no modifications. The Fock matrix and MO coefficients of the optimized excited reference are passed to the correlated calculation and all amplitudes (e.g. all singles and doubles in CCSD) are included. This procedure would not be of use for real applications because of the aforementioned tendency of variational collapse and instability of CC procedures on non-Aufbau references. Nevertheless, it provides useful insight in the few cases where the coupled cluster equations do converge. Such systems are few-atom molecules in a small basis, where there are no orbitals of the right energy to make the denominators small enough.

Figure 2.1 shows the basis set convergence of the CH_4 core ionization energies, as calculated with the Δ-based methods, with respect to increasing cardinality of the aug-cc-pCVXZ basis set. To compare, we use the experimental value of 290.83 $\pm$0.02 eV from Pireaux *et*
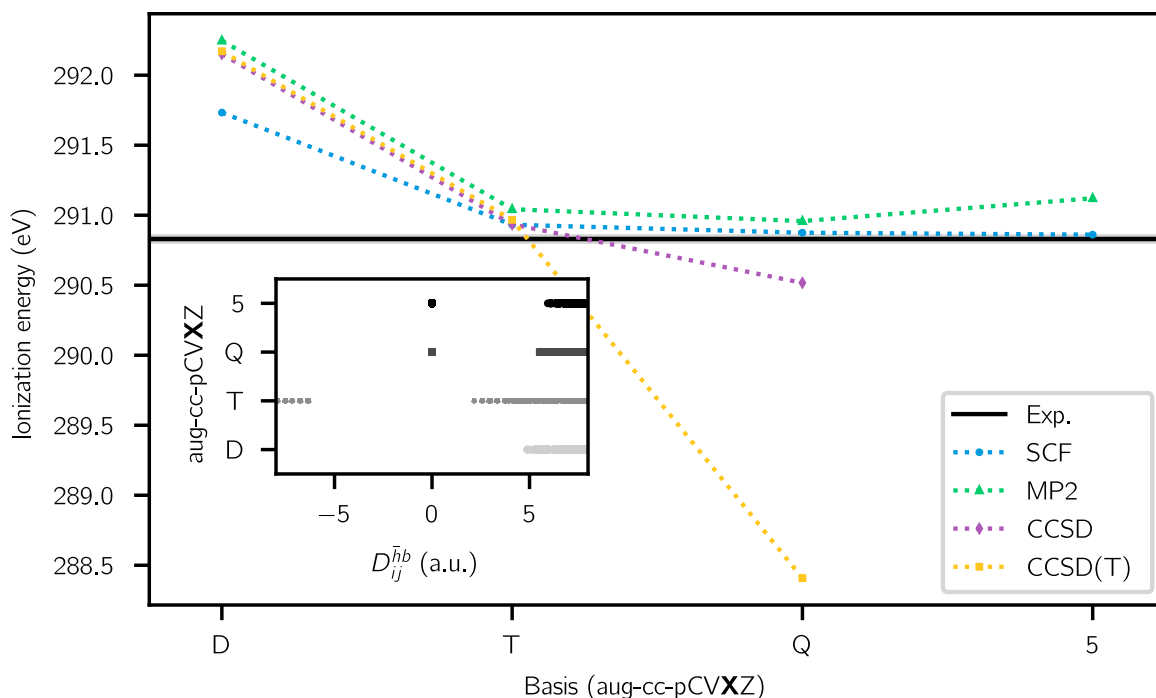
Figure 2.1: Core ionization energies for CH_4, calculated with Δ methods, as a function of basis set. Inset: energy denominators (Eq. 2.4) involving the virtual core orbital.

al., [209] but note that the compilation of Jolly *et al.* lists values from 290.8 - 291.0 eV. [210] The ΔSCF values converge quickly, with the TZ core ionization less than 0.1 eV above of the 5Z. The results for all the correlated Δ methods are within 0.1 eV of each other up until the QZ level, where they begin to diverge. At the 5Z level, the CCSD equations fail to converge and the ΔMP2 results break monotonicity. An analysis of the denominators associated with excitations into the core virtual (inset of Figure 2.1) reveals that, for all basis sets, there are positive denominators and, furthermore, that a close-to-zero denominator appears at the QZ level. Once the complexity of the molecule increases, the virtual space will begin to populate the problematic orbital energy range associated with near-zero denominators even when using small basis sets. Nonetheless, the proximity of the predicted ionizations with the experiment suggest a promising performance of the Δ-based methods if the irregularities caused by small denominators are addressed.
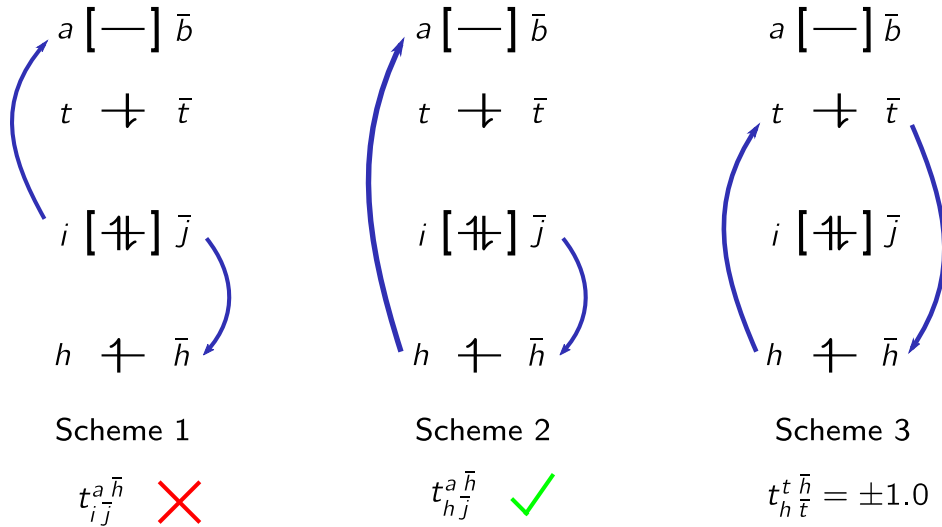
Figure 2.2: Visual summary of the schemes used to address the dangerous denominators.

2.4.1 Scheme 1: Deleting all amplitudes involving the core virtual

We make use of three schemes, presented visually in Figure 2.2, to address the numerical instabilities discussed previously. The first, which we refer as Scheme 1 (S1), is that proposed by Zheng *et al.*, and employed by Lee and Head-Gordon. [191, 192] This scheme simply excludes the core virtual from the correlation treatment. Additionally, we chose to exclude singles amplitudes that excite the occupied core electron.

$$\text{if}(a = \bar{h} \text{ or } i = h) \quad a_i^a, \, t_i^a = 0$$
$$\text{if}(a = \bar{h} \text{ or } b = \bar{h}) \quad a_{ij}^{ab}, \, t_{ij}^{ab} = 0$$
$$\text{if}(a = \bar{h} \text{ or } b = \bar{h} \text{ or } c = \bar{h}) \quad t_{ijk}^{abc}(c) = 0$$

Under these conditions, the ill-behaved amplitudes are removed by design. However, by excluding amplitudes that involve the core virtual, we are also excluding part of the correlation between the remaining core electron and valence electrons, as will be elaborated on the next Scheme. The de-excitation amplitudes in the Lambda equations, solved to obtain CC properties like $\langle S^2 \rangle$, are treated in a completely analogous way. Under these constraints, the Lambda equations converged to yield to similar $\langle S^2 \rangle$ values than without them, but at a much accelerated pace.

2.4.2 Scheme 2: Half-occupied core with zero spin-complement amplitude

To incorporate some of the correlation missing in S1, Scheme 2 (S2) allows for the double substitutions involving the core virtual, $\bar{h}$, that also promote the occupied electron in the same core orbital, h. These were found to be the leading amplitudes for some of the well-behaved calculations with no constrains at all. S2 is pleasing in that, even though core substitutions are involved, they are all associated with configurations that retain a core occupancy of 1.

$$\text{if}(a = \bar{h} \text{ or } i = h) \quad a_i^a,\ t_i^a = 0$$
$$\text{if}(a = \bar{h} \text{ or } b = \bar{h})$$
$$\text{if}(i \neq h \text{ or } j \neq h) \quad a_{ij}^{ab},\ t_{ij}^{ab} = 0$$
$$\text{if}(a = \bar{h} \text{ or } b = \bar{h} \text{ or } c = \bar{h})$$
$$\text{if}(i \neq h \text{ or } j \neq h \text{ or } k \neq h) \quad t_{ij k}^{abc}(c) = 0$$

As for S1, the CC de-excitation amplitudes are treated in a completely analogous way. For the core excitations using a spin-symmetry broken reference, we found that allowing for the double substitution that generates the spin complement of the reference $\hat{\alpha}_{h\bar{t}}^{t\bar{h}}$, with t being the target particle state, leads the CC iterations to converge towards the (lower energy) triplet excited state and resulting $\langle S^2 \rangle$ values that deviate significantly from 1. Therefore, an additional constraint was placed these calculations: the amplitude associated with the aforementioned excitation excitation is also set to zero. This helped ensure that the $\langle S^2 \rangle$ value of the CCSD wave function remained close to 1, signifying that it is a mixed spin configuration. As with S1, the spin contamination is removed by evaluating the singlet energy via Yamaguchi's AP expression.

2.4.3 Scheme 3: Half-occupied core with unit spin-complement amplitude

As a final scheme, and exclusively for the calculations on the mixed singlet state, we propose to incorporate all of the conditions of S2 but, instead of neglecting the double substitution amplitude $\hat{\alpha}_{h\bar{t}}^{t\bar{h}}$ associated with the spin complement of the reference, we set it to 1.0; we refer to this as Scheme 3 (S3). These conditions force the CC iterations to look for the pure singlet starting from the mixed reference. As previously, the exact same S3 conditions are imposed on the de-excitation amplitudes for the left eigenvectors of the similarity transformed Hamiltonian but we found this only accelerated the convergence of the Lambda equations. They were able to converge even without enforcing this condition, although at a slower pace. An attractive feature of S3, as will be elaborated on in the Results section, is that it bypasses the need for AP altogether because the resulting states have $\langle S^2 \rangle$ values relatively close to 0. S3 is, in fact, similar in spirit to the the bi-configurational MR-CC model proposed

by Oliphant and Adamowicz in 1991. [211] However S3 is dramatically simpler because additional triple and quadruple excitations that are necessary in MR-CC (in order to account for the single and double excitations on top of the "secondary reference") are omitted here.

The amplitude of the spin complement can also be set to -1.0 to access the $M_s = 0$ triplet. This allows us to asses the reliability of S3 by comparing its calculated triplet, $M_S = 0$ numbers against the $M_s = \pm 1$ triplet numbers obtained via S2. In the absence of spin-orbit coupling or external magnetic fields, the $M_s = 1$ and $M_s = 0$ triplet states should be degenerate, so any differences reflect the failures of S3 with respect to S2. Naturally, one source of error will be the fact that, in S3, the correlation methods treat each individual configuration of the CSF unequally.

2.5   Results and discussion

Figure 2.3 presents a visual summary of the statistical performance of ROKS(HF), ROKS(SCAN), fc-CVS-EOM-CCSD-EE, and the correlated Δ methods (Schemes S1, S2 and S3) in predicting K-shell core excitations relative to experimental values for 21 transitions. The the raw values for ROKS, fC-CVS-EOM-CCSD, and the best-performing ΔCCSD scheme are shown in Table 2.1 as calculated by ROKS(HF), ROKS(SCAN), fc-CVS-EOM-CCSD-EE, and the correlated Δ methods (Schemes S1, S2 and S3). Analogous information for K-shell core ionizations are presented in Figure 2.4 and Table 2.2 The raw numbers for all the correlated Δ methods methods, as well as the basis set convergence study for all methods, is provided in the Supplementary Information of the paper.

2.5.1   Excitations via ROKS and fc-CVS-EOM-CCSD

Before discussing the correlated methods, it is worth revisiting the ROKS results with the Hartree-Fock functional. For the excitations considered, ROKS(HF) achieves a mean absolute error (MAE) and root-mean squared-error (RMSE) of 0.43 and 0.52 eV. All of the excitations involving carbon and nitrogen, and the O 1s - σ^*/Rydberg transitions are overestimated. All of the fluorine and neon excitations , and the O 1s - π^* transitions are underestimated. This element-dependent error distribution with respect to experiment leads to a relatively small mean signed error (MSE) of 0.18 eV. Using ROKS with the standard SCAN functional, [212] the best-performing functional according to a recent study, [146] reduces the MAE, MSE, and RMSE to an impressive 0.16 eV, -0.08, and 0.19 eV. How well can CC methods compete with these results?

With a MAE and RMSE of 0.34 and 0.41 eV, fc-CVS-EOM-CCSD aproach cannot match ROKS(SCAN) and it scarcely outperforms the simple ROKS(HF). fc-CVS-EOM-CCSD tends to underestimate the excitations out of carbon, with an overestimation of 0.34 eV for the $\mathbf{CH_3OH}$ 1s $\longrightarrow$ 3s transition being the only serious exception. All other excitations are overestimated except for the N_2 1s $\longrightarrow \pi^*$ and Be 1s $\longrightarrow$ 2p excitations, which are underestimated by 0.25 and 0.68 eV. The latter might be a failure of the fc-CVS model.
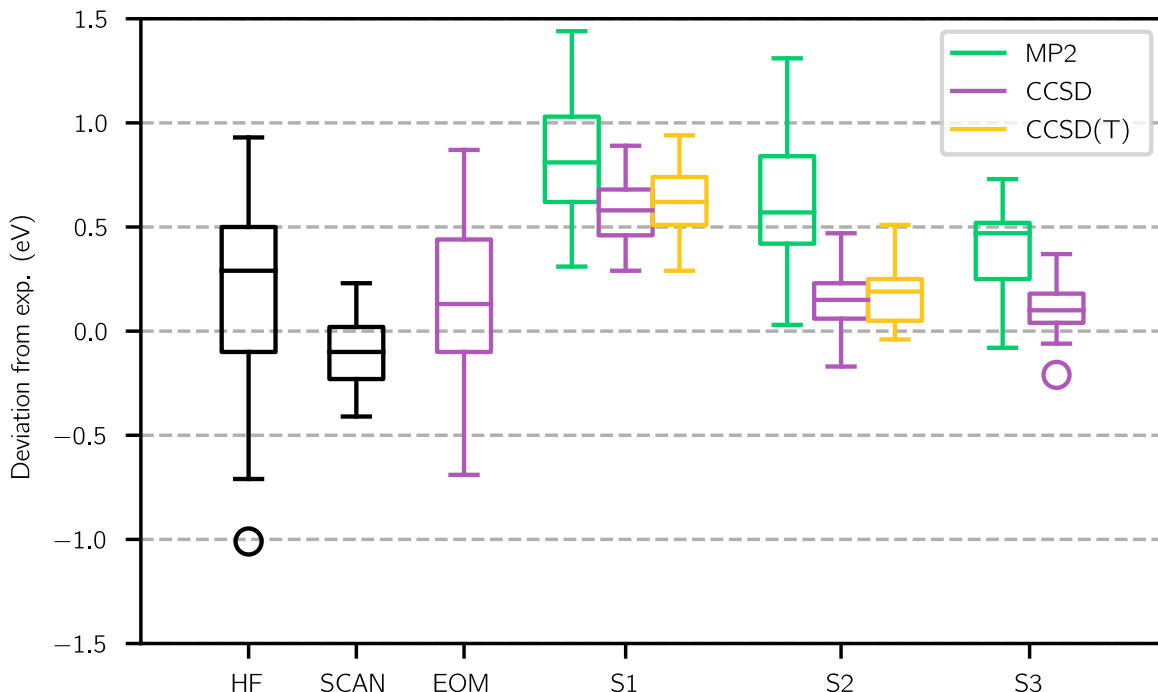
Figure 2.3: Statistical performance on the accuracy of different methods for predicting the 21 core excitations shown in Table 2.1.

2.5.2 Excitations via correlated Δ methods under S1, S2, and S3

Lets now turn our attention to the correlated Δ methods. Addressing the offending denominators, either by eliminating all excitations into the core virtual (S1) or including only those that retain a core occupancy of 1 (S2 and S3) resulted in well-behaved, monotonically convergent CC calculations in all cases. Furthermore, for Schemes S1 and S2, the MP2, CCSD, and CCSD(T) correlation energies of the excited states, and the calculated excitation energies seem to converge monotonically towards a well defined BSL.

Figure 2.3 and the SI reveal that the correlated calculations under Scheme 1 always overestimate the excitation energy. ΔMP2(S1), ΔCCSD(S1), and ΔCCSD(T)(S1) achieve MAEs of 0.82, 0.58, 0.63 eV, and RMSEs of 0.88, 0.60, 0.65 eV. ΔCCSD(S1) attenuates the most severe failures of ΔMP2(S1) - where it overestimates experiment by more than 1 eV: $H_2\mathbf{C}O$ 1s $\longrightarrow \pi^*$, $H\mathbf{C}N$ 1s $\longrightarrow \pi^*$, $HC\mathbf{N}$ 1s $\longrightarrow \pi^*$, N_2 1s $\longrightarrow \pi^*$, and F_2 1s $\longrightarrow \sigma^*$. These are all cases where ΔMP2(S1) changes the ROKS(HF) results the most - in all cases for worse - with F_2 having the largest change in magnitude, at 2.3 eV. ΔCCSD(T)(S1), more

often than not, seems to very slightly increase the error against experiment when compared to ΔCCSD(S1). Including correlation via S1, either via MP2, CCSD, or CCSD(T) only decreases the calculated values relative to ΔHF in roughly half the cases. The MSEs for all the correlated methods under S1 are identical to their MAEs, which is is consistent with a systematic overestimation of the excitation energies or, conversely, an under-correlation of the excited states. Since the results are expected to be well near the BSL, and the perturbative triples correction changes the CCSD results by a small amount, we attribute this to the configurations excluded from the correlation treatment for the sake of proper convergence.

As proposed in the previous section, not all configurations involving excitations into the core virtual need to be excluded for a safe convergence of the CC procedure. Figure 2.3 shows that including some of the missing configurations via scheme S2 indeed reduces the error relative to S1. ΔMP2(S2), ΔCCSD(S2), and ΔCCSD(T)(S2) achieves MAEs of 0.62, 0.18, and 0.20 eV, and RMSEs of 0.69, 0.22 and 0.25 eV. A small systematic overestimation remains, as suggested by MSEs of 0.61, 0.16, and 0.20 eV. Two relevant statistical observations are that ΔMP2(S2) still fails to offer an improvement over ROKS(HF), and that the (T) correction slightly worsens the ΔCCSD results. We note how the well-behaved excitations involving the core account for roughly 0.4 eV of the calculated excitation energy, as measured by the statistical differences between ΔCCSD(S1) and ΔCCSD(S2). This is in agreement with the findings of Zheng *et. al* and emphasizes that, if quantitative agreement is desired, a CVS scheme like S1 is inadequate. [191]

2.5.3 ΔCCSD triplet excitations: S2 ($M_s = 1$) vs. S3 ($M_s = 0$)

Before discussing the performance of S3 in predicting excitation energies, we make some other relevant remarks on the scheme. The de-excitation amplitudes usually converged without any modifications to yield a CCSD $\langle S^2 \rangle$ close to 0 (or 2, if the triplet state was being targeted). Naturally, it often takes many iterations for these amplitudes to respond to the large excitation amplitude in T_2. Imposing the condition analogous to S3 for the de-excitation amplitudes accelerated the convergence, never taking more than 35 iterations without DIIS for the cases that we studied. As is noted in the SI, a residual deviation from an $\langle S^2 \rangle$ value of 0 remained for all calculations. The largest of these deviations was for the C_2H_2 1s $\longrightarrow \pi^*$ state with an $\langle S^2 \rangle$ of 0.069, the average being 0.033. We suspect that this might be due to the missing excitations described in the discussion of S3.

The spin-forbidden excitations into the triplet $M_s = 0$ manifold were calculated with ΔCCSD(S3) by forcing the amplitude of the spin complement of the reference to be -1.0; they are listed in SI. We compared these against the triplet $M_s = 1$ excitation energies as calculated by ΔCCSD(S2). The largest deviation was of 0.09 eV for the H_2CO 1s $\longrightarrow \pi^*$ state, the average being 0.04 eV. The $M_s = 0$ triplet excitations were higher than the $M_s = 1$ results for all but one case, Be 1s $\longrightarrow$ 2p, where the difference is -0.01 eV. This is also consistent with the idea that for the $M_s = 0$ triplets, as for the singlets, we are undercorrelating the excited state due to missing excitations. An undercorrelation is not

present for the $M_s = 1$ triplet because, aside from any spatial symmetry breaking, this is purely a SR situation that S2 should be able to address. The triplet numbers, as calculated by ΔCCSD(S2), match fairly well with the two experimental numbers that we found for these spin-forbidden transitions: 114.3 eV for Be 1s $\longrightarrow$ 2p and 400.12 eV for N_2 1s $\longrightarrow \pi^*$. [213, 214] ΔCCSD(S2) predicts them to be 114.37 eV and 400.24 eV, respectively. The average energy difference between the singlet and triplet excited states for the set of molecules studied here, as calculated by ΔCCSD(S3), is 0.44 eV. Some cases worthy of notice are Be 1s $\longrightarrow$ 2p, where the splitting is 1.16 eV, and **CO** 1s $\longrightarrow \pi^*$, with the largest splitting of all: 1.42 eV. Interestingly, the splitting for C**O** 1s $\longrightarrow \pi^*$ is only 0.34 eV. Another case of relevance are the two Rydberg excitations Ne 1s $\longrightarrow$ 3s and Ne 1s $\longrightarrow$ 3p with the smallest splittings accross the data set: 0.06 eV and 0.05 eV.

2.5.4 Comparison of singlet excitations accross all methods

In Table 2.1, we present the calculated excitation energies of the singlet excited states for the most successful scheme, ΔCCSD(S3), against ROKS(HF), ROKS(SCAN), and fc-CVS-EOM-CCSD. All the statistics provided are compared against their most recent and / or accurate experimental values. The per-molecule results for the remaining schemes are listed in the SI. Overall, ΔCCSD(S3) achieves an MAE and RMSE of 0.14 and 0.18 eV. The most challenging excitation for this method is H_2**CO** 1s $\longrightarrow \pi^*$, with an overestimation of 0.37 eV from the experimental value of 287.98 eV by Remmers *et al.* [216]. A small systematic overestimation remains, as suggested by a MSE of 0.12 eV. The only excitation that ΔCCSD(S3) significantly underestimates is **CO** 1s $\longrightarrow \pi^*$, which is below Sodhi and Brion's result of 534.21 ± 0.09 eV by 0.21 eV. [218]

A recent study that is closely-related to our approach is the application of a direct two-determinant (TD) CCSD protocol to study core excited states. [195, 197] This procedure follows the ΔCC framework through orbital-optimizing a core excited configuration, constructing a CSF, and carrying out TD-CCSD on top of it. To address the dangerous denominators, an equivalent of our Scheme 2 is employed. [199] It is shown that TD-CCSD results have a comparable accuracy to the ΔCCSD results reported here, with a MAE of 0.10 eV and RMSE of 0.11 eV against the Coriani implementation of CVS-EOM-EE-CCSDT for the three lowest lying core excitations of HCN, CO, NH_3, and H_2O. The ΔCC approaches presented in our work have the advantage of halving the number of amplitudes as compared to the bi-configurational TD-CCSD, by virtue of employing pure SR formalism. Furthermore, employing the Scheme of choice to accelerate the convergence of the Lambda equations enables calculations of excited state properties such as gradients and $\langle S^2 \rangle$.

2.5.5 Core ionizations: ΔSCF vs. fc-CVS-EOM-IP-CCSD

Lets turn our attention to the performance of the calculated core ionizations, visualized in Figure 2.4. The experimental values used as a reference are the ones given by Jolly et al., [210] unless a more recent study was found. ΔSCF(HF) has a MSE, MAE, and RMSE of

Table 2.1: BSL estimate of the core excitation energies predicted by ROKS with the Hartree-Fock and SCAN functionals, ΔCCSD, and fc-CVS-EOM-CCSD compared against their most recent experimental values.

| Transition	HF	SCAN	ΔCCSD	EOM-CCSD	Exp.	Unc.	Ref.
Be 1s - 2p	115.37	115.34	115.53	114.79	115.47	-	[213]
C_2H_4 1s - π^*	285.27	284.70	284.77	284.68	284.68	0.1	[215]
H$_2$**C**O 1s - π^*	286.42	285.74	285.96	285.62	285.59	-	[216]
C_2H_2 1s - π^*	286.40	285.67	285.84	285.55	285.9	0.1	[215]
H**C**N 1s - π^*	286.98	286.35	286.51	286.07	286.37	-	[217]
CO 1s - π^*	288.05	286.99	287.46	286.71	287.40	0.02	[218]
CH$_3$OH 1s - 3s	288.91	288.18	288.34	288.26	287.98	-	[219]
CH$_4$ 1s - 3p(t_2)	288.38	287.96	288.02	287.90	288.00	0.2	[82]
HC**N** 1s - π^*	400.00	399.60	399.80	399.74	399.7	-	[217]
NH$_3$ 1s - 3s	400.97	400.42	400.63	400.82	400.66	0.2	[82]
N$_2$ 1s - π^*	401.18	400.80	401.02	400.63	400.88	0.02	[218]
NH$_3$ 1s - 3p(e)	402.62	402.18	402.41	402.46	402.33	0.2	[82]
H$_2$C**O** 1s - π^*	530.67	530.83	530.86	531.26	530.82	-	[216]
H$_2$O 1s - 3s	534.15	533.84	534.14	534.44	534.0	0.2	[82]
CH$_3$**O**H 1s - 3s	534.16	533.98	534.24	534.64	534.12	-	[219]
C**O** 1s - π^*	533.68	533.97	534.00	534.50	534.21	0.09	[218]
H$_2$O 1s - 3p (b$_2$)	536.03	535.65	536.08	536.21	535.9	0.2	[82]
F$_2$ 1s - σ^*	681.19	682.43	682.41	683.07	682.2	0.1	[220]
HF 1s - σ^*	687.31	687.44	687.76	688.05	687.4	0.2	[220]
Ne 1s - 3s	864.75	865.18	865.37	865.54	865.1	0.1	[220]
Ne 1s - 3p	866.58	866.96	867.30	867.40	867.29	-	[221]
MSE	0.15	-0.09	0.12	0.11			
MAE	0.43	0.15	0.14	0.34			
RMSE	0.52	0.19	0.18	0.41			
MAX	1.01	0.41	0.37	0.87			
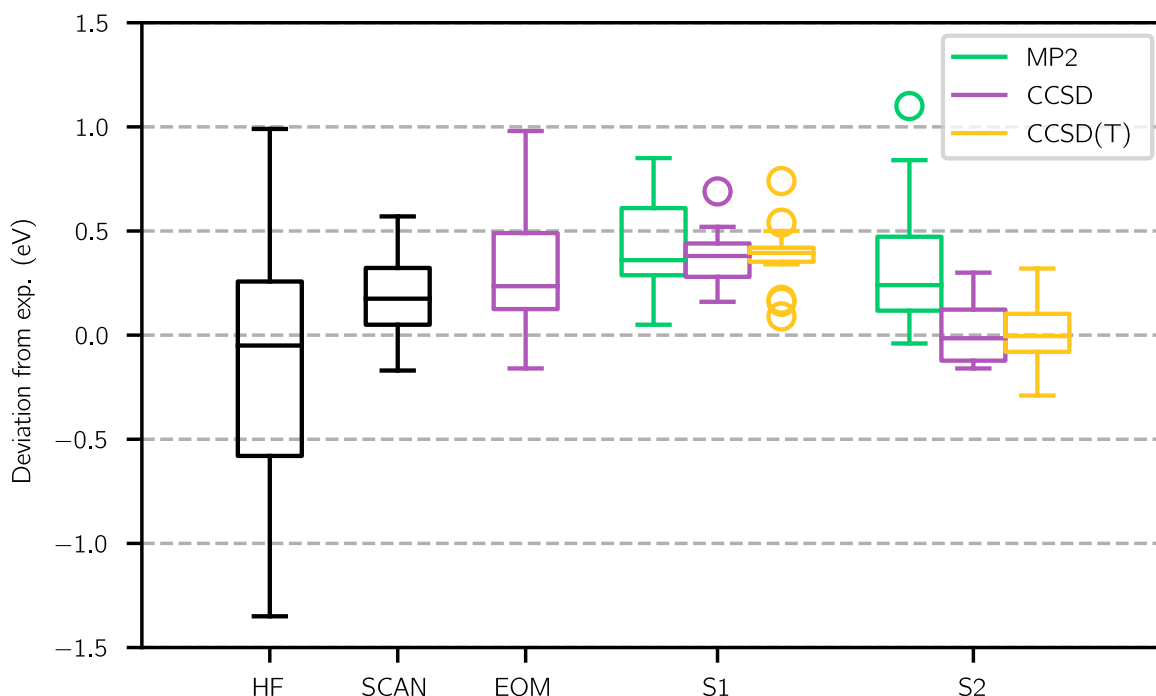
Figure 2.4: Statistical performance on the accuracy of different methods for predicting the
21 core ionizations shown in Table 2.2

-0.15, 0.45, and 0.58 eV, respectively. The two most challenging cases for ΔSCF(HF) in
the ionization data set, CO and F_2, are the only cases with an error greater than 1 eV.
ΔSCF(SCAN) reduces the ΔSCF(HF) errors by more than a factor of two, with an MAE
and RMSE of 0.21 and 0.25 eV. In contrast to excitations, all ionizations except two, F_2 and
Ne, are overestimated with ΔSCF(SCAN), resulting in an MSE similar to its MAE: 0.18 eV.
The most challenging case for ΔSCF(SCAN) is Be, over estimated by 0.51 eV. Somewhat
surprisingly ΔSCF(HF) predicts the Be experimental ionization perfectly.

The performance of ΔSCF(HF) against the much more sophisticated fc-CVS-EOM-IP-
CCSD is once again remarkable, with the MAE and RMSE of the latter being 0.35 and 0.45
eV. Elaborating on the previous discussion on the specific details of the CVS implementation,
we note that these fc-CVS-EOM-IP-CCSD errors are roughly five times smaller than those
reported by Liu et al. for the Coriani-style CVS-EOM-IP-CCSD. [138]

2.5.6 Ionizations via correlated Δ methods under S1 and S2

In contrast to excitations, the correlated Δ methods using the S1 model manage to slightly improve upon ΔHF for ionization. ΔMP2(S1) increases the HF ionization energy in almost all cases, and over 1 eV in several of them: $H_2C\mathbf{O}$, $CH_3\mathbf{OH}$, $C\mathbf{O}$, HF, F_2, and Ne. The only case where ΔMP2(S1) decreases the ionization predicted by ΔHF is $\mathbf{C}$O, which is also the second most challenging case for ΔHF, right after F_2. The problematic Be is overestimated by 0.81 eV by ΔMP2(S1). Once again, ΔCCSD(S1) alleviates the worst cases in ΔMP2(S1). $\mathbf{C}$O is anomalous in that this is the only case where ΔCCSD(S1) significantly worsens the ΔMP2(S1) result, and also the only one where the (T) seems to significantly improve the result, correcting the ΔCCSD(S1) result by 0.17 eV. Overall, the S1 methods result in MAEs and RMSEs of 0.42, 0.37, 0.38 eV and 0.49, 0.39, 0.41 eV for MP2, CCSD, and CCSD(T). As Lubijic noted in their study, ΔMP2(S1) seldom warrants the additional cost over ΔSCF and neither extending to CCSD or CCSD(T) seems to improve the results to an extent that justifies their cost. [190]

A consistent overestimation of the core ionization energies, as suggested by the MSEs being equal to the MAEs for all the S1 correlated methods, once again hints at the correlation lost to S1 having a measurable consequence. Indeed, the improvement in calculated core ionization energies provided by the correlated methods under model S2, relative to S1, is even more dramatic than it is for the excitations. In contrast with ΔMP2(S1), ΔMP2(S2) manages to somewhat improve the statistics from ΔHF, bringing down the MAE and RMSE to 0.33 and 0.44 eV. S2 improves the S1 results for MP2 in almost all cases, the only significant exception being Be, where ΔMP2(S2) performs the worst: an overestimation of 1.1 eV. As with S1, ΔCCSD(S2) alleviates the failures of ΔMP2(S2) (significantly for Be) and brings the MAE and RMSE down to 0.12 and 0.15 eV. ΔCCSD(T) slightly worsens the statistics by bringing the MAE and RMSE to 0.13 and 0.17 eV. The RMSE for ΔCCSD(S2) is more than 2.5 times smaller than for fc-CVS-EOM-IP-CCSD.

Table 2.2 compares the ΔSCF(HF), ΔSCF(SCAN), ΔCCSD(S2), and fc-CVS-EOM-IP-CCSD core ionizations against experiment; the results for ΔCCSD(S2) presented here are comparable to those in Table 5 of Zheng *et al.*. [191] The differences can be associated with the different basis sets used and the way we are treating the correlation associated with the core virtual. Whereas in their study, they make estimates to the correlation missing due to freezing the core orbital completely (S1) by carrying out unconstrained calculations with denominator thresholds, S2 recovers it by a well-defined protocol.

2.6 Conclusions

We have studied the use of core-hole orbital-optimized references in single-reference correlated methods to describe core excited and core ionized states of 18 small closed-shell organic molecules, and compared them against two of the most successful approaches so far: ROKS(SCAN) and fc-EOM-EOM-CC. The use of three different schemes (S1, S2, S3) to

Table 2.2: BSL estimate of the core ionization energies predicted by ΔSCF with the Hartree-Fock and SCAN functionals, ΔCCSD, and fc-CVS-EOM-IP-CCSD compared against their most recent experimental values.

Molecule	HF	SCAN	ΔCCSD(S2)	EOM-CCSD	Exp.	Unc.	Ref.
Be	123.35	123.92	123.65	123.49	123.35	-	[213]
C_2H_4	290.71	290.92	290.72	290.95	290.88	-	[210]
CH_4	290.86	290.92	290.69	290.68	290.83	-	[210]
C_2H_2	291.39	291.47	291.21	291.26	291.14	-	[210]
CH₃OH	292.63	292.63	292.44	292.52	292.3	0.2	[222]
H**C**N	293.76	293.68	293.43	293.34	293.50	-	[210]
H₂**C**O	294.91	294.75	294.50	294.70	294.35	-	[216]
CO	297.23	296.58	296.47	296.43	296.24	-	[210]
NH_3	405.48	405.70	405.51	405.77	405.52	-	[210]
HC**N**	406.74	406.96	406.78	407.10	406.8	-	[210]
N_2	410.21	410.15	409.99	409.89	409.9	-	[210]
CH₃**O**H	538.43	539.08	538.90	539.64	539.06	0.2	[222]
H₂C**O**	538.51	539.47	539.29	540.28	539.30	-	[216]
H_2O	539.49	539.96	539.82	540.29	539.92	-	[210]
C**O**	541.79	542.65	542.43	543.10	542.57	-	[210]
HF	693.62	694.30	694.25	694.80	694.0	-	[210]
F_2	695.36	696.54	696.58	697.58	696.71	-	[210]
Ne	869.54	870.21	870.31	870.49	870.33	-	[221]
MSE	-0.15	0.18	0.02	0.31			
MAE	0.45	0.21	0.13	0.35			
RMSE	0.58	0.25	0.17	0.45			
MAX	1.35	0.57	0.33	0.98			

address the convergence problems of the CC equations, and the spin contamination of the excited states, were employed. S1 excludes all amplitudes involving the half-occupied core orbital associated with the excitation or ionization. S2 allows for the ones that retain a core occupancy of 1. S3, exclusively for CCSD on core excitations, fixes the T_2 amplitude associated with the spin compliment of a spin symmetry-broken core excited reference to ± 1.0, thereby ensuring the proper reference CSF is present in the cluster expansion. As evidenced by the energetic difference between the singlet and the triplet core excited states, addressing the spin contamination associated with using a symmetry broken reference is essential for quantitative studies using the correlated Δ methods unless Rydberg states are being targeted.

To compare with experimental core excitations and ionizations requires careful attention to basis set convergence, which we have addressed by using the aug-cc-pCVXZ basis set for

heavy atoms (n = T, Q, with extrapolation), and aug-cc-pVDZ for hydrogen. With this protocol, ΔCCSD(S3) performs the best among the correlated Δ methods for core excitations, reaching an MAE and RMSE of 0.14 and 0.18 eV for CCSD. These statistics are on par with the most successful orbital-optimized DFT approach, ROKS(SCAN). ΔCCSD(S2) follows closely behind, with an MAE and RMSE of 0.18 and 0.22 eV. As such, ΔCCSD with either S2 or S3 roughly halves the errors of fc-CVS-EOM-EE-CCSD. A similar situation takes place for ionizations, where S2 in conjunction with CCSD performs the best, by achieving a MAE and RMSE of 0.13 and 0.17 eV, respectively. ΔCCSD(S2) reduces the fc-CVS-EOM-IP-CCSD error by more than a factor of 2.5 and outperforms ΔSCF(SCAN), which has an MAE and RMSE of 0.21 and 0.25 eV.

The use of a CVS scheme like S1 for the correlated Δ methods is discouraged, if quantitative agreement is sought after. Furthermore, as has previously been concluded by others, we cannot recommend the use of ΔMP2 for the prediction of core excitations or ionizations. [190] In the future, it may be interesting to explore whether regularization or further orbital re-optimization can address the limitations of ΔMP2. [25, 192] Finally, we note that the use of the perturbative (T) triples correction with the best scheme that allows for it, S2, does not seem to offer a significant improvement over CCSD.

There are additional sources for the disagreement with regards to experimental values. Difficulties in measuring X-ray spectra often result in slightly different experimental values from different sources (see Ref. [221] or [210], for example) which are often on the order of the errors observed here. We have made our best effort to obtain the most recent and reliable information available at the moment. Additionally, physical effects lacking in our model may also contribute to a disagreement with the experiment. There are two such effects that we expect to be of relevance. The first is the fact that we are treating core excited states as formally bound, whereas in reality they are resonances coupling with the Auger continuum. [223] Said effect is expected to shift the energy of the resonance. The second is that we are computing vertical excitation energies - a more complete model would incorporate vibronic effects. [85, 86, 224, 225]

Despite its shortcomings, the main tool for routine calculation of XAS is TDDFT. Furthermore, due to the recent advances in LR-DFT-based theory, [114, 173, 176] the efficient implementations of ΔSCF methods, [163] and specialized basis sets, [202] techniques based on mean field approaches will likely remain the workhorses for the calculation of core spectra. Nonetheless, considering an accuracy of less than 0.2 eV attained by ΔCCSD, we expect this to be a promising way to provide benchmark theory-based core ionization numbers. With recent refinements on the formalism for singlet excited states, we expect state-specific coupled-cluster methods to also provide excellent benchmark numbers for core excitations. [226]

The challenges to making ΔCCSD a practical method for the calculation of excitation energies, as can now be done with EOM-CCSD, is largely implementational. Specialized and efficient amplitude windowing algorithms to carry out the particular ΔCC scheme and a robust workflow that allows for the ΔCCSD calculation on a number of states of interest (which can be carried out in parallel) could eventually lead to routine ΔCCSD calcula-

tions for transition energies. Furthermore, the question of compact and efficient basis sets for these orbital-optimized, wave-function-based correlated calculations deserves future attention. New developments for the calculation of transition properties, such as oscillator strengths, within the ΔCCSD framework are still needed in order to make this approach an attractive alternative to conventional CC methods for excited states.

Chapter 3

Generalization of one-center non orthogonal configuration interaction singles to open shell singlet reference states

3.1 Introduction

Time-resolved near-edge X-ray absorption (TR-NEXAS) experiments aim to track the ensuing dynamics of molecular systems after a perturbation with light by monitoring the NEXAS (equivalently referred to as NEXAFS, XANES, or often simply XAS) features of the species involved as a function of time. With advances in synchrotron slicing techniques and the advent of free-electron lasers, the time resolution of modern TR-NEXAS experiments is well into the femtosecond regime. [3, 227] Furthermore, improvements in high-harmonic generation have brought extreme UV and soft X-ray femtosecond pulses in the water window (270 - 550 eV) to table-top laser equipment. [228, 229] With element and site specificity, as well as strong sensitivity to the electronic environment of the species being probed, the TR-NEXAS experiments enabled by these new technologies have already provided fundamental insight into the role of dark singlet and triplet states in the electronic relaxation of organic molecules and directly tracked the nuclear motion of small molecules post strong-field ionization. [89–94]

Before the development of ultra-fast techniques, the experimental focus of NEXAS was on characterizing stable molecules in their ground state. Accordingly, the development of electronic structure methods to aid in the assignment of NEXAS features focused predominantly on closed-shell systems as the reference state. A comprehensive review of the variety of methods available for computing NEXAS of closed shell systems is beyond the scope of this article, but refer to Section 1.3 for a review of some of those that are relevant to this work.

3.1.1 Challenges to core excited states of radicals

Relative to the development in electronic structure methods for computing the NEXAS of closed shell systems, the theoretical modeling of the NEXAS of open-shell radicals remains in an exploratory stage due to a collection of challenges. First, common to both closed-shell and radical systems, an electronic structure model must account for core-hole relaxation to obtain reasonable accuracy. Second, open-shell systems are often multi-configurational: aside from the case where the initial state is a high-spin open-shell, such as a doublet with a single open shell or an $M_s = \pm 1$ triplet with two open shells, multiple configurations are necessary for the proper description of a spin-pure initial state reference. Generating excited states out of an open-shell initial state compounds the challenge of ensuring spin purity. Third, electronic structure NEXAS calculations on short-lived radical species rely on molecular dynamics (MD) simulations for nuclear geometries that properly represent the evolution of the system, often involving a wide range of configurations, imposing the need for efficient generation of the spectra.

The underdevelopment of methods for core spectroscopies on open-shell methods, coupled with the experimental advances and exciting prospects for unconvering fundamental chemical phenomena, have fueled a rapidly-growing body of work. While a variety of theories, such as 1C-NOCIS and EA-CIS, [170–172] EOM-CCSD, [139] ΔSCF and ROKS, [149, 150] and a number of TDDFT-based formalisms, [230] have been extended to treat one-electron open-shell (1eOS) doublets, high-spin 2eOS triplets, or arbitrary high-spin systems, we focus on the developments for 2eOS singlets and highlight relevant ideas from other open-shell cases when useful.

3.1.2 Core transitions into singly-occupied molecular orbitals

Figure 3.1 provides a visual guideline to the types of excited states relevant in UV-pump X-ray-probe TR-NEXAS experiments on a molecule with a closed-shell ground state. A visible or UV pump causes one photon absorption that promotes the molecule to an optically allowed valence excited state where an electron in **o**ccupied orbital (**o**) has been promoted to a **t**arget virtual orbital (**t**):

$$|^1\Phi_o^t\rangle = (2)^{-1/2}\left(|\Phi_o^t\rangle + |\Phi_{\bar{o}}^{\bar{t}}\rangle\right) \tag{3.1}$$

Three kinds of core excited states can be conceived out of a 2eOS singlet valence excited state. When the **c**ore electron (**c**) re-pairs either with the now singly-occupied molecular orbital (SOMO) in the occupied space, or with the newly created particle SOMO, we can obtain two singlet CSFs:

$$|^1\Phi_{oc}^{to}\rangle = |^1\Phi_c^t\rangle = (2)^{-1/2}\left(|\Phi_c^t\rangle + |\Phi_{\bar{c}}^{\bar{t}}\rangle\right) \tag{3.2}$$

$$|^1\Phi_{oc}^{tt}\rangle = (2)^{-1/2}\left(|\Phi_{\bar{o}c}^{\bar{t}t}\rangle + |\Phi_{o\bar{c}}^{t\bar{t}}\rangle\right) \tag{3.3}$$
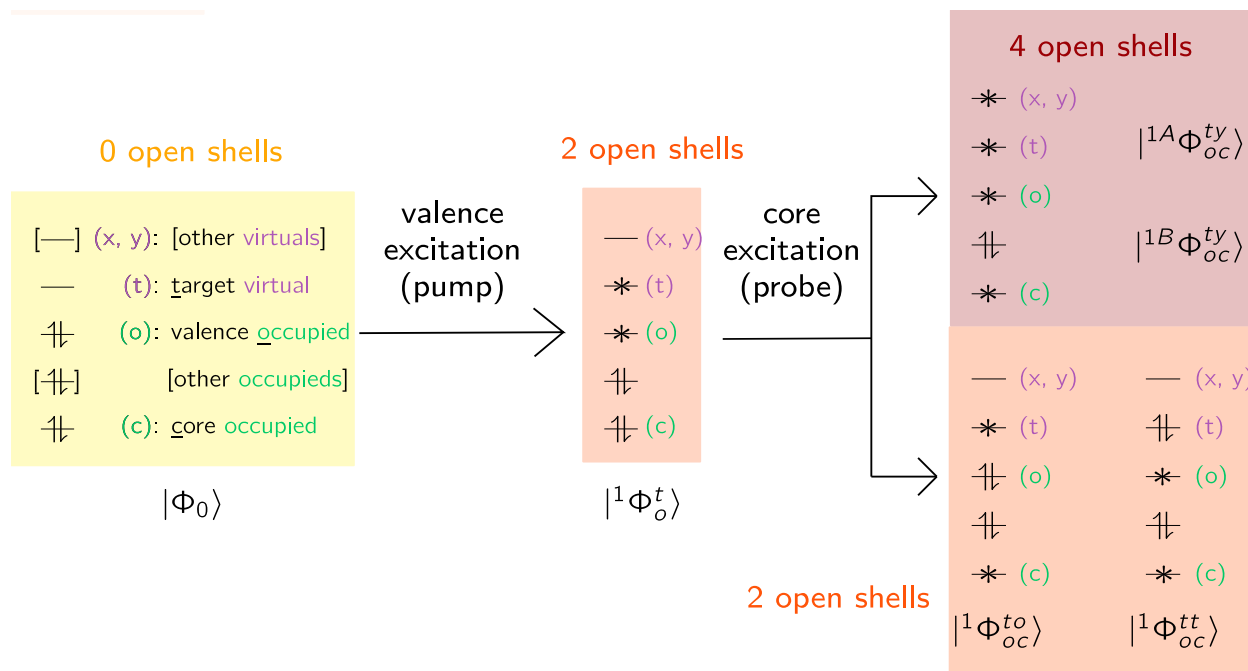
Figure 3.1: Electronic states resulting from valence-core pump-probe excitations.

Transitions into SOMOs are spectral features characteristic of open-shell systems in general. The c → SOMO(o) transition, associated with Eq. 3.2, is special in that it is usually well-separated to lower energy from the NEXAS transitions in the ground state and allows for a clear detection of a valence excited state in the TR-NEXAS when bright. [89, 90] Nonetheless, its ability to distinguish between valence excited states that share the same hole levels is limited. [91] Either OO-DFT or response theories, such as TDDFT, ADC, and EOM-CC , can reliably calculate the c → SOMO(o) transition as an energy difference between the valence excited state and the pump-probe core excited state, since the later can be reached from the ground state by a single excitation. [231]

The c → SOMO(t) state associated with Eq. 3.3, on the other hand, is beyond traditional TDDFT because it is a double excitation out of the closed-shell reference. As it is well-established that EOM-CC requires truncation beyond doubles to properly correlate doubly-excited states, the $|^1\Phi^{tt}_{oc}\rangle$ excited state is beyond EOM-CCSD despite the fact that the dominant configuration exists within the associated Fock space. [232] For similar reasons, it is possible that the modification of ADC to describe 2eOS singlets proposed by Ruberti *et al.* also fails to capture the $|^1\Phi^{tt}_{oc}\rangle$ excited state when truncated at second order. [233, 234] Regardless, the former was used by Neville and coworkers to calculate the NEXAS (as well as the extended X-ray absorption fine structure - EXAFS) of a few excited states of ethylene, ammonia, carbon dioxide, and water at their Frank-Condon geometries. [235] Subsequently, they employed the model to predict and study the TR-NEXAS of ethylene in detail, demon-

strating the sensitivity of NEXAS to nuclear motion and electronic character, even if solely judged by the c $\to$ SOMO transitions. [236, 237] Very recently, the aforementioned ADC methodology was also used to predict the TR-NEXAS of pyrazine after excitation into its B_{2u} $\pi \to \pi^*$ state at the nitrogen K-edge. [238] Two DFT-based methods proposed in the last two years explore unusual response approaches to capture both c $\to$ SOMO transitions. The first is the hole-hole Tamm-Dancoff Approximated (hh-TDA) DFT, which shows promise in its efficiency yet still suffers from a lack of orbital relaxation and the difficulties associated with converging a doubly-electron attached reference. [239] In contrast, mixed-reference spin-flip (MR-SF)-TDDFT employs both $M_s = 1$ and $M_s = -1$ open-shell core-excited triplet references, relaxed accordingly, to access both the valence excited singlet state and the c $\to$ SOMO configurations via spin-flip operations. [176]

3.1.3   Core transitions into fully-vacant molecular orbitals

Continuing with the higher energy one-electron core excitations, we arrive at those involving a core-excitation into a fully-vacant virtual orbital. These types of transitions, resulting in a four-electron open-shell (4eOS) excited state, are beyond traditional TDDFT, ADC(2), EOM-CCSD, or hh-TDA DFT. MR-SF-TDDFT can only partially describe the 4eOS states, since some of the configurations necessary are still not accessible with the theory. There are six 4eOS $M_S = 0$ configurations with SOMOs in o, c, t, and an arbitrary virtual orbital y.

$$|\Phi_{oc}^{ty}\rangle,\ |\Phi_{\bar{o}\bar{c}}^{\bar{t}\bar{y}}\rangle,\ |\Phi_{\bar{o}c}^{\bar{t}y}\rangle,\ |\Phi_{o\bar{c}}^{t\bar{y}}\rangle,\ |\Phi_{\bar{o}c}^{\bar{y}t}\rangle,\ |\Phi_{o\bar{c}}^{y\bar{t}}\rangle \tag{3.4}$$

As demonstrated in Section 3.1 of the Supporting Information (SI), diagonalizing the S^2 operator in the basis of the six 4eOS, $M_S = 0$ configurations yields two linearly independent singlet, three triplet, and one quintet CSFs. [10, 40]

$$|^{1_A}\Phi_{oc}^{ty}\rangle = (12)^{-1/2}\left(2\,|\Phi_{oc}^{ty}\rangle + 2\,|\Phi_{\bar{o}\bar{c}}^{\bar{t}\bar{y}}\rangle + |\Phi_{\bar{o}c}^{\bar{t}y}\rangle + |\Phi_{o\bar{c}}^{t\bar{y}}\rangle - |\Phi_{\bar{o}c}^{\bar{y}t}\rangle - |\Phi_{o\bar{c}}^{y\bar{t}}\rangle\right) \tag{3.5}$$

$$|^{1_B}\Phi_{oc}^{ty}\rangle = (2)^{-1}\left(|\Phi_{\bar{o}c}^{\bar{t}y}\rangle + |\Phi_{o\bar{c}}^{t\bar{y}}\rangle + |\Phi_{\bar{o}c}^{\bar{y}t}\rangle + |\Phi_{o\bar{c}}^{y\bar{t}}\rangle\right) \tag{3.6}$$

$$|^{3_C}\Phi_{oc}^{ty}\rangle = (2)^{-1/2}\left(|\Phi_{oc}^{ty}\rangle - |\Phi_{\bar{o}\bar{c}}^{\bar{t}\bar{y}}\rangle\right) \tag{3.7}$$

$$|^{3_D}\Phi_{oc}^{ty}\rangle = (2)^{-1/2}\left(|\Phi_{\bar{o}c}^{\bar{t}y}\rangle - |\Phi_{o\bar{c}}^{t\bar{y}}\rangle\right) \tag{3.8}$$

$$|^{3_E}\Phi_{oc}^{ty}\rangle = (2)^{-1/2}\left(|\Phi_{\bar{o}c}^{\bar{y}t}\rangle - |\Phi_{o\bar{c}}^{y\bar{t}}\rangle\right) \tag{3.9}$$

$$|^{5_F}\Phi_{oc}^{ty}\rangle = (6)^{-1/2}\left(|\Phi_{oc}^{ty}\rangle + |\Phi_{\bar{o}\bar{c}}^{\bar{t}\bar{y}}\rangle - |\Phi_{\bar{o}c}^{\bar{t}y}\rangle - |\Phi_{o\bar{c}}^{t\bar{y}}\rangle + |\Phi_{\bar{o}c}^{\bar{y}t}\rangle + |\Phi_{o\bar{c}}^{y\bar{t}}\rangle\right) \tag{3.10}$$

A common strategy to generate an excited-state NEXAS - including the 4eOS states - is to employ TDDFT or EOM-CCSD for the core excitation on top of an optimized non-Aufbau configuration representing the initial valence excited state. [90, 91, 231, 240] ΔSCF - the aforementioned procedure to capture the valence excited states - employs a single unrestricted configuration to describe the valence excited state, rendering it severely spin-contaminated

for open-shell singlets. A response theory that also disregards spin-symmetry on top of it carries over and exacerbates the spin-contamination on the excited states, artificially shifting the predicted energies and possibly predicting spurious bright excitations (Section 1 of the SI). Incorporating the developments in alternative approaches to reconcile economic standard response theories with open-shell references while addressing spin purity, such as the spin-adapted (s)-TDDFT is an exciting prospect. [241–243]

While using ΔSCF in conjunction with response theories spin-contaminates the resulting pump-probe states, employing ΔSCF to target both the initial valence excited states and the final pump-probe core excited states allows for spin-purification procedures, such as the approximate spin-projection (AP), to address the deficiency of employing a single configuration to describe low-spin open-shell excited states. [144] The AP is straightforward for 2eOS singlets and the spin-purification can be made rigorous by employing a single set of restricted open-shell orbitals for the AP procedure, like in the ROKS approach, instead of optimizing unrestricted configurations separately for the low-spin and high-spin orbitals. Extending the spin-recoupling schemes to 3eOS and 4eOS systems opens the door for the full NEXAS calculation beyond the c $\rightarrow$ SOMO transitions of doublet radicals and 2eOS singlets using AP-ΔSCF. [149, 150, 244] In a similar vein, Zhao and co-workers have used multi-state (MS)-DFT to produce spin-pure core excited states out of open-shell radicals, with the attractive feature of addressing the correlation double-counting inherent to employing DFT orbitals in CI-like formalism. [154] Recently, a state-specific approach relying on CI employing core-relaxed orbitals to provide partially sin-complete 4eOS excited states was recently used by Garner and Neuscamman to explore the effect of spin coupling of the core excited states and their sensitivity to nuclear geometry for photochemical ring-opening of furanone. [245] The down-side to state-specific methods like AP-ΔSCF, ROKS, MS-DFT, or the CI method reported in Ref. [245] arises from their inconvenience. In the case of AP-ΔSCF or MS-DFT, a number of unrestricted configurations (two, four, and eight for 2eOS singlets, 3eOS doublets, and 4eOS singlets, respectively) must be optimized independently to calculate spin-pure excited states. Difficult convergence and asserting that the configurations indeed correspond to the same set of spatial orbitals presents a challenge to automation procedures, making this approach cumbersome for systems of moderate size and for generating a large number of spectra for different nuclear configurations. While progress in this area has been accomplished for 2eOS core excited states via AP-ΔSCF, [163] no such automation procedure has been designed for 3eOS and 4eOS excited states.

Real-time (RT)-TDDFT provides an alternative to response theories for the simulation of excited state absorption; the interested reader is directed to Section 4.3 of the review by Li *et al.*[246] Leveraging said capacity, it has recently been employed to simulate TR-NEXAS. [247, 248] In theory, RT-TDDFT is able to describe the whole spectrum of excited states but it suffers from practical considerations. First and foremost, and in common with standard frequency-domain TDDFT, the adiabatic approximation results in an inability to describe core-hole relaxation and causes RT-TDDFT to incur in large shifts to match experimental profiles in the X-ray regime unless using specially-tailored functionals. Furthemore, employing "CVS-like" schemes to decouple the core resonances from the continuum presents

additional challenges. [249] Second, the need to explicitly propagate the electronic density in time makes TR-NEXAS simulations in the hundred-femtosecond time-scales of nuclear motion intractable at the moment, with the ones reported at the moment extending to only a few attoseconds. [247, 248]

The last three excited-state methodologies we are aware of to calculate pump-probe core excited states, including the 4eOS states absent from the majority of the theories previously described, involve explicit use of multi-reference (MR) formalism. The first is a MR configuration-interaction (CI) procedure employing DFT orbitals, namely DFT-MRCI, [250] employed by Seidu and coworkers to simulate the excited-state NEXAS of 1, 3-butadiene. [251] The second involves carrying out a complete active space (CAS)SCF calculation, or its dynamically-correlated PT2 variant, in conjunction with restricted-active space (RAS) techniques to ensure the proper occupancy of the relevant orbitals: one for the core orbital, one for the valence hole level, and one for the valence particle level. RASSCF / RASPT2 have been used to theoretically probe the relaxation pathways of malonaldehyde with XAS, [252] and to predict the NEXAS signatures of the $\pi \rightarrow \pi^*$ and $N_{LP} \rightarrow \pi^*$ states of azobenzene. [253] The third is a generalized-active space driven similarity renormalization group technique, [254] recently employed to explore the NEXAS signatures of three aromatic diradicals. [255] A multi-reference adaptation to ADC can provide the NEXAS of singlet excited states in theory but that capacity has not been explored yet. [256] Their own problems aside, namely electron correlation double-counting for DFT-MRCI, the intruder state phenomena for RAS-SCF / RAS-PT2, and the computational cost that often comes with MR methodologies, these are promising approaches capable of simulating both the complex non-adiabatic dynamics relevant for relaxation after excitation, and the full NEXAS spectra of the resulting species.

3.1.4 Objective statement and summary of the work

The near-absence of electronic structure models capable of providing properly core-hole-relaxed, spin-pure pump-probe core excited states in an efficient manner motivates the present work. We present a new method which generalizes EA-CIS and the related NOCIS methods to describe K-shell core excitations from a 2eOS singlet excited state of a molecule with a closed shell ground state. [109, 170, 172] The paper begins by presenting a variety of reference orbitals with the relaxation appropriate for the description of core excited states out of a valence excited state. Relying on these orbitals, the ansatz for the spin-adapted generalization of the one-center (1C)-NOCIS to 2eOS singlets is then provided. By virtue of working with configuration state functions (CSFs), the method avoids shifts in the energy due to spin-contamination and, perhaps more importantly, ensures that transitions are bright only when they ought to be. While 1C-NOCIS 2eOS relies on state-specific methods to generate different sets of orbitals, it provides the full core spectrum given a target valence excited state and the core edge of interest. The approach proposed here has an interesting connection with the recent line of work by Kossoski and Loos on ΔCI, hierarchy-CI, and seniority-CI. [257, 258] We elaborate on the subject in Sections 3.2.1 and 3.2.2
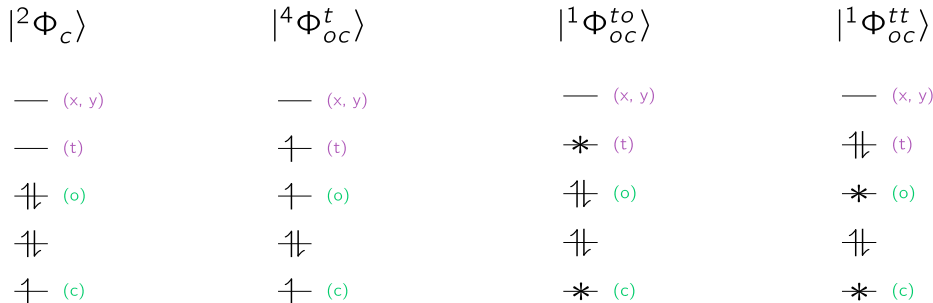
Figure 3.2: Different choice of reference orbitals considered for the construction of the pump-probe core excited states.

With the 1C-NOCIS 2eOS protocol established, calculations on water and thymine as test systems are carried out to assess the general NEXAS features of valence excited states relative to each other and to the ground state. The ability of 1C-NOCIS to efficiently simulate spectra is showcased by a simulation of the TR-NEXAS of acetylacetone at the carbon K-edge after excitation into its lowest $\pi \rightarrow \pi^*$ state, based on a sequence of snapshot structures from molecular dynamics simulations.

3.2   Theory

3.2.1   Reference orbitals for the pump-probe excited states

We begin by introducing the reference orbitals that we will use for the description of the pump-probe core excited states, visualized in MO diagrams in Figure 3.2. Appropriate core-hole orbital relaxation is key for an accurate description of core excited states, whether for closed-shell references or open-shell species. However, open-shell singlet excited states feature an additional challenge arising from the partially-filled orbitals. Ideally, the reference orbitals would also take into consideration the orbital relaxation arising from the valence excitation which, albeit smaller than the core-hole relaxation, may still be relevant.

ROHF doublet core ion orbitals ($|^2\Phi_c\rangle$)

One attractive candidate set of orbitals comes from ROHF-optimized $M_S = \frac{1}{2}$ doublet core-ionized references, $|^2\Phi_c\rangle$, for which the presence of the core hole confer the orbitals the appropriate contraction. The deficiency of this choice lies in the poor description of valence particle levels afforded by the canonical virtual orbitals of the core-ionized reference. To address the need for a well-defined target particle level **t** we rotate the doublet core ion virtual space into the natural transition orbital (NTO) basis of a regular 1C-NOCIS calculation for

the ground state which provide core particle levels in resemblance with valence particle levels. This is similar in essence to what was done by Hait et al. to evaluate the ROKS energies from a STEX calculation employing DFT orbitals. [173] The target virtual is simply chosen as the one with the highest overlap with the particle level of the valence excited state. Even with a well-defined particle level, this choice disregards the relaxation in the orbitals due to the valence excitation. A second shortcoming presents itself when the valence particle level has a fortuitously-large overlap with a core particle level of different character, or when there is simply no clear connection between the valence particle level and any core particle level. We conceive this situation taking place for high-lying valence excited states with loosely-defined particle levels, such as Rydberg states in complex molecules.

ROHF quartet core ion orbitals ($|^4\Phi_{oc}^t\rangle$)

An alternative choice of orbitals could come from optimizing the high-spin $M_s = 3\,/\,2$ quartet core ion associated with the valence excited state, $|^4\Phi_{oc}^t\rangle$, via ROHF. The advantage of this choice comes in that the valence particle level **t** is explicitly optimized and, furthermore, the rest of the orbitals are optimized in presence of an electron in said orbital. In other words, this set of orbitals incorporate relaxation due to the valence particle level as well as the core hole. The disadvantage comes in that the spatial description of the orbitals of a singlet state may differ significantly from those optimized for a quartet reference, a phenomena that MR SF-TDDFT likely also suffers from when trying to describe singlet excited states with orbitals optimized for a triplet state.

ROKS core excitation orbitals ($|^1\Phi_c^t\rangle$)

A third choice is to construct the CSFs from the $|^1\Phi_c^t\rangle$ configuration optimized via ROKS. Like the $|^4\Phi_{oc}^t\rangle$ orbitals, this choice is appealing because it produces a set of orbitals relaxed in the presence of a core hole **c** and a particle level **t**, refined by the orbital optimization procedure. Unlike the $|^4\Phi_{oc}^t\rangle$ orbitals, the spatial description of the $|^1\Phi_c^t\rangle$ configuration is optimized within the correct multiplicity. Note that the ROKS procedure is susceptible to nonphysical mixing between the two open-shells when they possess the same spatial symmetry that must be addressed to avoid overly-intense oscillator strengths (Section S5.2 of the SI). [161] A final choice of orbitals we considered come from ROKS optimization of the $|^1\Phi_{oc}^{tt}\rangle$ configuration. These orbitals take into consideration relaxation due to the valence hole **o** as well as the core hole **c**. We disregard this choice from here on, as the ROKS optimization of this configuration is challenging in practice.

Alternative choices and practical considerations

An appealing alternative choice of orbitals, unavailable to us at the moment, is evident: ROKS optimization of the 3eOS doublet core ion, where the three open-shells lie in the **c**, **o** and **t** orbitals. This configuration is perfect in that it accounts for relaxation due to all the holes and particles present in the pump-probe excited states associated with a particular

valence excited state. While two linearly independent 3eOS doublet states exist (Section S3.1 of the SI), the most sensible choice for describing singlet states would be the 3eOS doublet genealogically-related with, say, the $|^1\Phi_c^t\rangle$ singlet. [40]

As demonstrated in the work of Kossoski and Loos, state-specific CI calculations are sensitive to the underlying choice of orbitals. [258] We summarize our take on the advantages and disadvantages of the different sets of orbitals available to us for the construction of the pump-probe states, to be explained in Section 3.2.2. The $|^1\Phi_c^t\rangle$ orbitals optimize the target particle level $\mathbf{t}$ and the core hole $\mathbf{c}$ within a state of singlet multiplicity, and are thus the preferred orbitals when convergence of the ROKS SCF equations is possible. The $|^2\Phi_c\rangle$ orbitals are often easier to obtain because converging them requires ROHF as opposed to ROKS but they do not optimize the target particle level. Finally, the $|^4\Phi_{oc}^t\rangle$ orbitals do optimize the particle level but within the quartet multiplicity and, as a result, these orbitals are probably more appropriate to describe triplet core excited states than singlet core excited states. The most rigorous way to test these hypotheses would be to compare against high level benchmark from theory or experimental data of high resolution. As will be elaborated on in Section 3.4.4, none are available at the moment.

For the purposes of obtaining the sets of orbitals (one for each core orbital of interest) 1C-NOCIS 2eOS relies on either ΔSCF or ROKS. In other words, this step of the process uses state-specific methodology and is non-black box: it requires the user to specify the valence excited state and the core orbitals of interest. To automate this process as much as possible, our implementation uses the valence excited-state orbitals, which define the $\mathbf{o}$ and $\mathbf{t}$ orbitals, as a guess for the core-hole orbital-relaxation procedure via ROHF or ROKS. Defining the core orbitals of interest is never a problem since those of each element are well separated energetically and can be singled out according to the edge desired.

While converging high-energy core-excited-state solutions is often difficult, there has been tremendous progress on the past decades. The techniques we employ for that purpose include the maximum-overlap methods, the state-targeted energy projection scheme, and square-gradient minimization, which provides a reliable suite of techniques to stabilize most core excited states. [140–143] Even still, converging the required orbitals remains one of the challenges to 1C-NOCIS 2eOS but progress on this subject is active and we expect advances to further streamline this step. [259] As a concluding remark, we mention that localization of the core orbitals (when the canonical ground-state orbitals delocalize over several atoms) prior to the SCF re-optimization in the presence of a core hole is crucial for improvements in accuracy. [169]

3.2.2 Ansatz for the 1C-NOCIS 2eOS wave function

With a variety of reference orbitals available, we proceed to introduce our model for the pump-probe excited states. For convenience, we introduce a set of basis functions for each core orbital of interest in the system built out of CSFs orthogonalized against the initial 2eOS state $|\Psi_i\rangle = |^1\Phi_o^t\rangle$. The initial state can obtained by either orbital-optimization via ROKS or constructed from a CIS wave function for the state of interest rotated into the

NTO basis and truncated to the dominant contributor. Note that this limits the 1C-NOCIS 2eOS model to valence excited states that do not feature genuine configuration-interaction and can be described with a single CSF.

$$|^1\tilde{\Phi}^t_c\rangle = (1 - P_i)\,|^1\Phi^t_c\rangle \tag{3.11}$$

$$|^N\tilde{\Phi}^{tx}_{oc}\rangle = (1 - P_i)\,|^N\Phi^{tx}_{oc}\rangle \tag{3.12}$$

In Eqs. 11 and 12, P_i represents a projector of the initial state, and in Eq. 12, the N labels the linearly independent states of a specific spin (Eqs 5 - 10). When speaking of determinant-based excitations, there are some doubly-excited configurations in Eqs. 3.11 and 3.12 with respect to either of the configurations in the valence excited state $|^1\Phi^t_o\rangle$ (the last two configurations in Eq. 3.4). However, a more natural description of the space spanned by Eqs. 3.11 and 3.12 is all the possible "singly-excited CSFs" for a given valence state $|^1\Phi^t_o\rangle$, restricted to excitations from a target core orbital in the spirit of the CVS scheme. Using the notion of CSF excitations, and under the CVS scheme, 1C-NOCIS corresponds to a ΔCIS method in the context of the Refs. [257] and [258], but employing different sets of orbitals for the initial state and the final states (hence the non-orthogonality). Perhaps because their scope is more broad, with a stronger focus on capturing the correlation of the reference state and only a minor component on a study of the excited states, ΔCIS was not studied in the work of Kossoski and Loos. On the other hand, our objective is very specialized: to describe the core excited states of a valence excited state.

To that end, once a choice of orbitals has been made from the candidates presented in the Section 3.2.1, the core-excited CSFs are constructed. The Hamiltonian in the basis of these projected-out CSFs takes the following form:

$$\tilde{H}^{MN}_{x,\,y} = \langle^M\Phi^{tx}_{oc}|\,(1 - P_i)H(1 - P_i)\,|^N\Phi^{ty}_{oc}\rangle \tag{3.13}$$

$$= \langle^M\Phi^{tx}_{oc}|\,H\,|^N\Phi^{ty}_{oc}\rangle + \langle^M\Phi^{tx}_{oc}|\,H\,|^1\Phi^t_o\rangle\,\langle^1\Phi^t_o|^N\Phi^{ty}_{oc}\rangle +$$
$$\langle^M\Phi^{tx}_{oc}|^1\Phi^t_o\rangle\,\langle^1\Phi^t_o|\,H\,|^N\Phi^{ty}_{oc}\rangle + \langle^M\Phi^{tx}_{oc}|^1\Phi^t_o\rangle\,\langle^1\Phi^t_o|\,H\,|^1\Phi^t_o\rangle\,\langle^1\Phi^t_o|^N\Phi^{ty}_{oc}\rangle \tag{3.14}$$

$$\tilde{\mathbf{H}}^{MN} = \mathbf{H}^{MN} + \mathbf{H}^M_{\text{N.O.}}(\mathbf{S}^N_{\text{N.O.}})^T + \mathbf{S}^M_{\text{N.O.}}(\mathbf{H}^N_{\text{N.O.}})^T + E_i \cdot \mathbf{S}^M_{\text{N.O.}}(\mathbf{S}^N_{\text{N.O.}})^T \tag{3.15}$$

The first term in Eq. 3.15 dominates by far. The remaining terms contribute small corrections due to the non-orthogonality between the valence excited state CSF and the final pump-probe CSFs arising from a different orbital basis. Naturally, the $\mathbf{S}_{\text{N.O.}}$ and $\mathbf{H}_{\text{N.O.}}$ matrices must be built with non-orthogonal configuration interaction techniques. For the 4eOS configurations, we label the Hamiltonian matrix elements with only the subscripts $\mathbf{x}$ and $\mathbf{y}$, denoting arbitrary virtual orbitals with respect to the closed-shell configuration, because this is the only running index; the indexes $\mathbf{o}$, $\mathbf{t}$, and $\mathbf{c}$ are fixed by the valence excited state being probed and the core orbital of interest. In other words, the dimensionality of the matrix $\tilde{H}^{MN}_{x,\,y}$ is, at most, $N_{\text{vir}} \times N_{\text{vir}}$.

Figure 3.3 provides a visual representation of the 1C-NOCIS 2eOS Hamiltonian for singlet pump-probe excited states. While the 2eOS configuration associated with the c $\rightarrow$ SOMO(o) transition must be explicitly included, the 2eOS configuration associated with the c $\rightarrow$
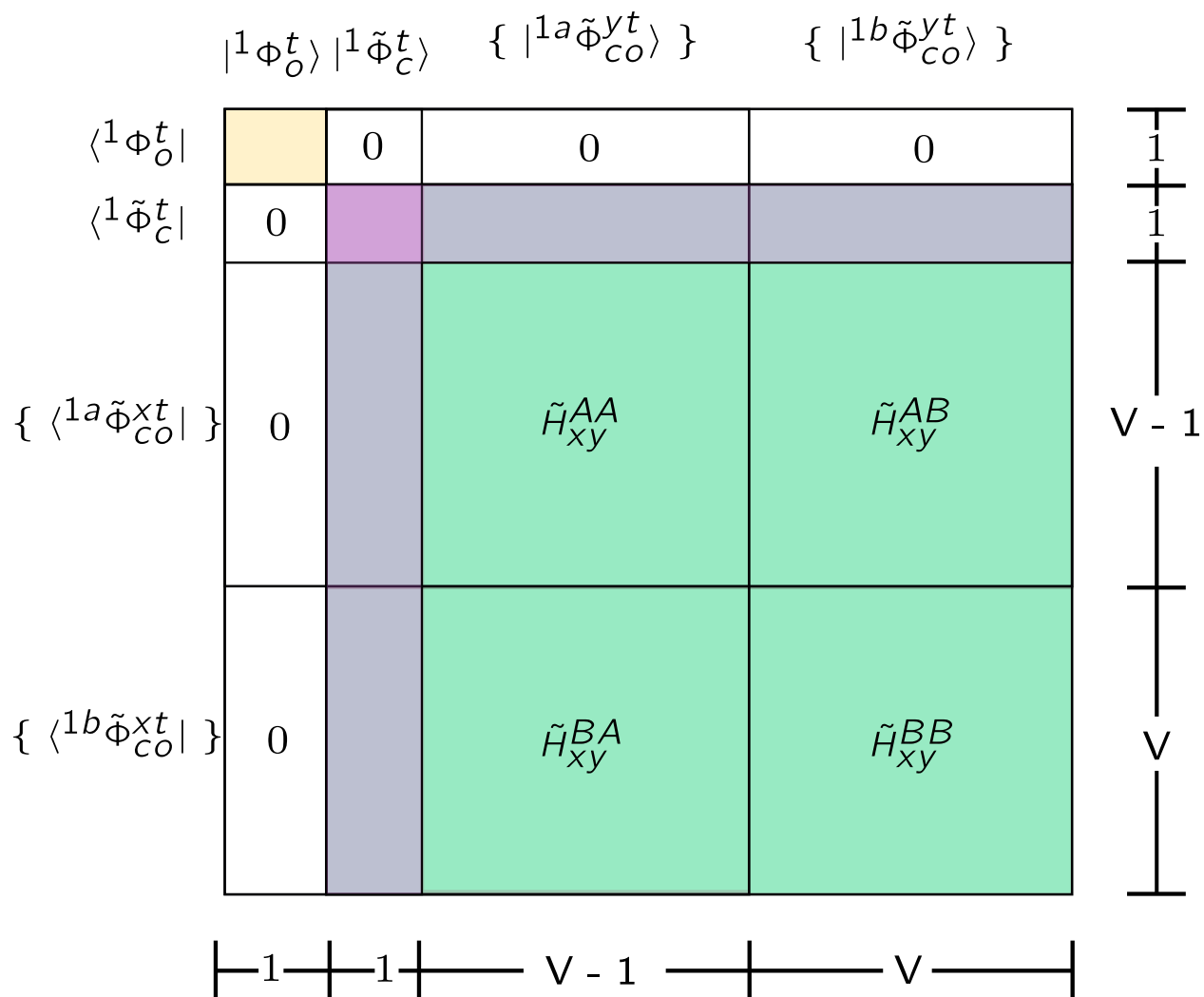
Figure 3.3: Visual representation of the 1C-NOCIS 2eOS Hamiltonian for singlet pump-probe excited states.

SOMO(t) transition is included as a special case of the 4eOS $|^{1B}\Phi_{oc}^{ty}\rangle$ configuration when y = t. On the other hand, the $|^{1A}\Phi_{oc}^{ty}\rangle$ configuration vanishes when y = t. In total, the dimensionality of the 1C-NOCIS 2eOS singlet Hamiltonian is $2N_{\text{vir}} \times 2N_{\text{vir}}$. Section 2 in the SI provides derivations of the standard non-relativistic electronic Hamlitonian in the basis of the relevant non-spin-adapted configurations via algebraic techniques. [60] Section 3.2 provides the Hamiltonian matrix elements in the basis of CSFs by taking the appropriate linear combinations of the non-spin adapted matrix elements dictated by the eigenfunctions of the S^2 operator, derived in Section 3.1; these matrix elements correspond to the first term in Eq 16. Section 4 provides simplifications due to spin-adaptation to $\mathbf{S}_{\text{N.O.}}$ and $\mathbf{H}_{\text{N.O.}}$. Solving the eigenvalue problem

$$\tilde{\mathbf{H}}\mathbf{C} = \tilde{\mathbf{S}}\mathbf{C}\mathbf{E} \tag{3.16}$$

where the matrix

$$\begin{aligned}
\tilde{S}_{x,y}^{MN} &= \langle^M\Phi_{oc}^{tx}|\,(1-P_i)(1-P_i)\,|^N\Phi_{oc}^{ty}\rangle \\
&= \langle^M\Phi_{oc}^{tx}|\,(1-P_i)\,|^N\Phi_{oc}^{ty}\rangle \\
&= \langle^M\Phi_{oc}^{tx}|^N\Phi_{oc}^{ty}\rangle + \langle^M\Phi_{oc}^{tx}|^1\Phi_o^t\rangle\,\langle^1\Phi_o^t|^N\Phi_{oc}^{ty}\rangle
\end{aligned} \tag{3.17}$$

$$\tilde{\mathbf{S}} = \mathbf{I} - \mathbf{S}_{\text{N.O.}}^M(\mathbf{S}_{\text{N.O.}}^N)^T \tag{3.18}$$

accounts for the deviation from orthogonality between the CSFs due to the projection of the initial state yields Hamiltonian eigenstates

$$|\Psi_f\rangle = b_c^t\,|^1\tilde{\Phi}_c^t\rangle + \sum_{M,y} b_x^\sigma\,|^M\tilde{\Phi}_{oc}^{ty}\rangle \tag{3.19}$$

$$|\Psi_f\rangle = (1-P_i)\left(b_c^t\,|^1\Phi_c^t\rangle + \sum_{M,x} b_x^M\,|^M\Phi_{oc}^{ty}\rangle\right) \tag{3.20}$$

with eigenvalues E_f. The excitation energy is simply the energy difference between the initial valence excited state and the final pump-probe states, $\Delta E_{if} = E_f - E_i$. Since this step of the process is diagonalization-based, it is completely black-box. Once the reference orbitals are defined, the whole core spectrum for the target valence excited state is generated. Oscillator strengths in the length gauge

$$f_{f\leftarrow i} = \frac{2}{3}\Delta E_{if}|\,\langle\Psi_i|\,\hat{\mu}\,|\Psi_f\rangle\,|^2$$

are the final ingredient to generate such a theoretical spectra. Central to this quantity is the transition dipole moment, $\langle\Psi_i|\,\hat{\mu}\,|\Psi_f\rangle$. While comparing the predicted oscillator strengths in the length gauge against the results in the velocity or mixed gauge could provide valuable information, we did not consider it in this study and leave it to future investigations.

3.3   Computational methods

1C-NOCIS 2eOS was implemented in a development version of QChem 6.1 and will be available in the next public release. [260] Calculations were carried out with different aug-pcX-n (n = 1 - 2) basis sets on the atoms associated with the K-edge of interest, with an aug-pcseg-1 basis on the remaining atoms. [202, 261] Specifically, the basis set combinations employed for each of the three systems considered is as follows:

- Water: aug-pcX-2 (O) and aug-pcseg-1 (H)

- Thymine: aug-pcX-2 on the atoms of the edge probed, with aug-pcseg-1 on all the other atoms.

- Acetylacetone: aug-pcX-1 (C) and aug-pcseg-1 (O, H)

Scalar relativistic effects are incorporated via the X2C model. [262] The geometry used for the calculations on water was optimized at the Hartree-Fock level of theory. The geometry used for the calculations on thymine, provided in the literature, [89] was optimized to the ground state at the CCSD(T) / aug-cc-pVDZ level. Two NAMD trajectories on acetylacetone for proof-of-concept spectral simulation were carried out using the augmented fewest-switches surface-hopping (AFSSH) algorithm. [263] They were initiated on the first $\pi \to \pi^*$ state from the ground state geometry, optimized at the Hartree-Fock level, and carried out with CIS with the aforementioned basis set combination for consistency of the dynamics with the NEXAS calculations.

3.4   Results

3.4.1   Preliminary excited-state NEXAS calculations

Before a discussion on the excited state spectra, we make an important note about the range of applicatibility of our 1C-NOCIS calculations, both for the ground state and the excited states. Without an appropriate treatment of the continuum of states beyond the K-edge ionization threshold - the region associated with the EXAFS - our 1C-NOCIS calculations are only appropriate for computing the bound core resonances associated with the NEXAS. As is evident from Figure S3, states beyond the ionization threshold are ill-behaved and we thus set the ionization threshold as a boundary to the reliability of our theoretical spectrum. The ionisation threshold for the closed-shell ground state is simply defined as the energy difference between the ground state and the optimized core-ionized reference - namely the ΔSCF ionization energy. Removing a core electron from a 2eOS valence excited state results in a 3eOS system, and thus there are three different ionization potentials for the valence

excited states. These are associated with the quartet and two doublet states.

$$|^{2G}\Phi^t_{oc}\rangle = (6)^{-1/2}\left(2\,|\Psi^t_{o\bar{c}}\rangle - |\Psi^t_{o\bar{c}}\rangle - |\Psi^t_{c\bar{o}}\rangle\right)$$

$$|^{2H}\Phi^t_{oc}\rangle = (2)^{-1/2}\left(|\Psi^t_{o\bar{c}}\rangle - |\Psi^t_{c\bar{o}}\rangle\right)$$

$$|^{4I}\Phi^t_{oc}\rangle = (3)^{-1/2}\left(|\Psi^t_{o\bar{c}}\rangle + |\Psi^t_{o\bar{c}}\rangle + |\Psi^t_{c\bar{o}}\rangle\right)$$

We establish the ionization thresholds for the valence excited states as the energy difference between the valence excited state and the energy of the 3eOS CSFs evaluated with whichever orbitals are employed for the 1C-NOCIS 2eOS procedure.

3.4.2   Water

The 1C-NOCIS NEXAS for the closed-shell ground state of water is in excellent agreement with experiment, requiring only a shift of -0.29 eV to align the strongest signal. Figure 3.4 diplays the 1C-NOCIS 2eOS NEXAS for nine different valence excited states of water, associated with the different combinations of $2p_x$ (b_1), $2p_z$ (a_1), and $2p_y$ (b_2) hole levels with 3s (a_1), $3p_y$ (b_2), and $3p_z$ (a_1) particle levels. The **o** and **t** orbitals of the core-hole-relaxed reference, serving as the visual indexes for the NEXAS of the different excited states in Figure 3.4, clearly resemble the particle and hole levels of the valence excited states of water. The c $\rightarrow$ SOMO(o) transitions dominate the spectrum of the excited states and are red-shifted from the ground state features by several eVs. The strong intensity of the $O_{1s} \rightarrow$ SOMO(o) transitions is due to the localization of both the 1s and 2p orbitals in the oxygen atom, allowing for a large transition dipole matrix element $\langle\phi_{O_{1s}}|\hat{\mu}|\phi_{O_{2p}}\rangle$. In contrast the overlap between the 1s orbital and the diffuse Rydberg orbitals is smaller and results in a lower transition intensity into said particle levels. This region of the spectrum, containing the transitions that are analogous to the transitions out of the closed-shell ground state, is richer than the latter due to the spin-splitting in the 4eOS core excited states (Figure S6).

A striking feature of the 4eOS states is a strong blue shift on the order of 5 - 10 eV for the high-energy features relative to the ground state transitions, with a general increase in the oscillator strength. While these states seem to be all beyond the ionization threshold of the ground state, a few of them are still within the core ionization threshold corresponding to the valence excited state itself. The two previous calculations we found on the K-edge ionization energy of the first two excited states of water at the Frank-Condon geometry differ drastically. The Neville and coworkers report an IP of roughly 530 eV using CVS-IP-ADC(2)-x results of, compared to the IP of the ground state at 540 eV. [235] Instead, the ΔSCF / CVS-EOM-IP-CCSD results of Moitra *et al.* predict the different doublet and quartet ionization potentials to lie at 544 - 549 eV. [264] The ionization thresholds computed in this work, in the range of 550 - 554 eV and visualized as vertical lines in Figure 3.4, are in closer agreement to those of Moitra and co-workers. While the 1C-NOCIS 2eOS results presented here, the ΔSCF / CVS-EOM-EE-CCSD calculations of Moitra *et al.*, and the CVS-ADC(2)-x work of Neville *et al.* agree on the strong c $\rightarrow$ SOMO(o) transition red
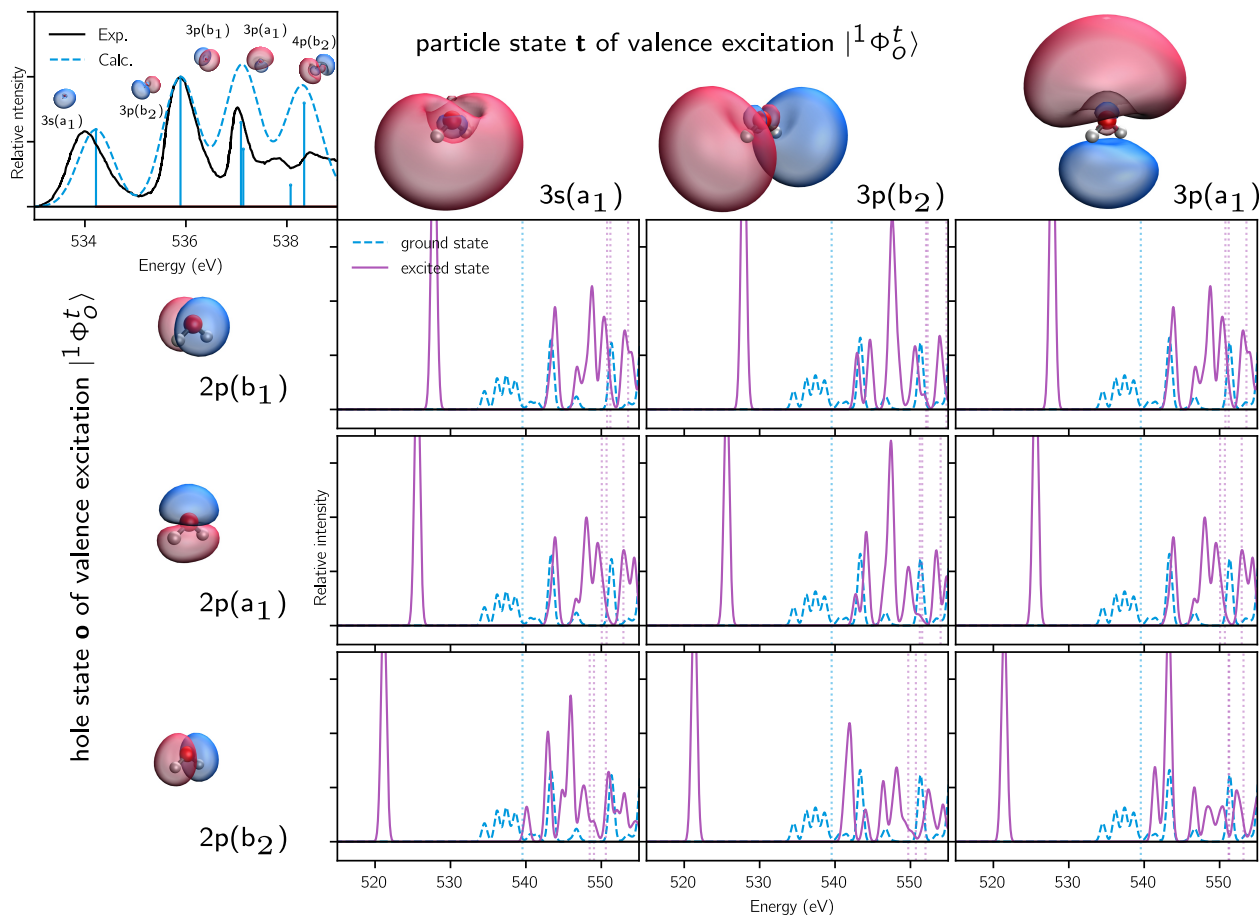
Figure 3.4: 1C-NOCIS NEXAS for nine different singlet valence excited states of water compared to the calculated ground state NEXAS. Top-left inset: 1C-NOCIS NEXAS of the ground state of water vs. experiment.

shifted from the ground state, the fate of the c $\rightarrow$ SOMO(t) 2eOS and the 4eOS states differs significantly. As the authors acknowledge, and as explained in the Background section, the high-energy transitions are beyond ΔSCF / CVS-EOM-EE-CCSD and are completely absent from the spectra presented in their work. CVS-ADC(2)-x, predicts these transitions to be well into the ionization continuum, beyond 560 eV.

We close this section by briefly commenting on the impact of the choice of orbitals on the excited state NEXAS, and elaborate on Section 5.3 of the SI. The spectral profile generated the ROHF-optimized $|^2\Phi_c\rangle$ on the 1C-NOCIS NTO basis and ROKS-optimized $|^1\Phi_c^t\rangle$ orbitals are indistinguishable from each other (Figure S5). On the other hand, the spectra resulting from the ROHF-optimized $|^4\Phi_{oc}^t\rangle$ orbitals differs significantly from those generated with the other two candidate orbitals. Since no experimental data exists yet for these states, the
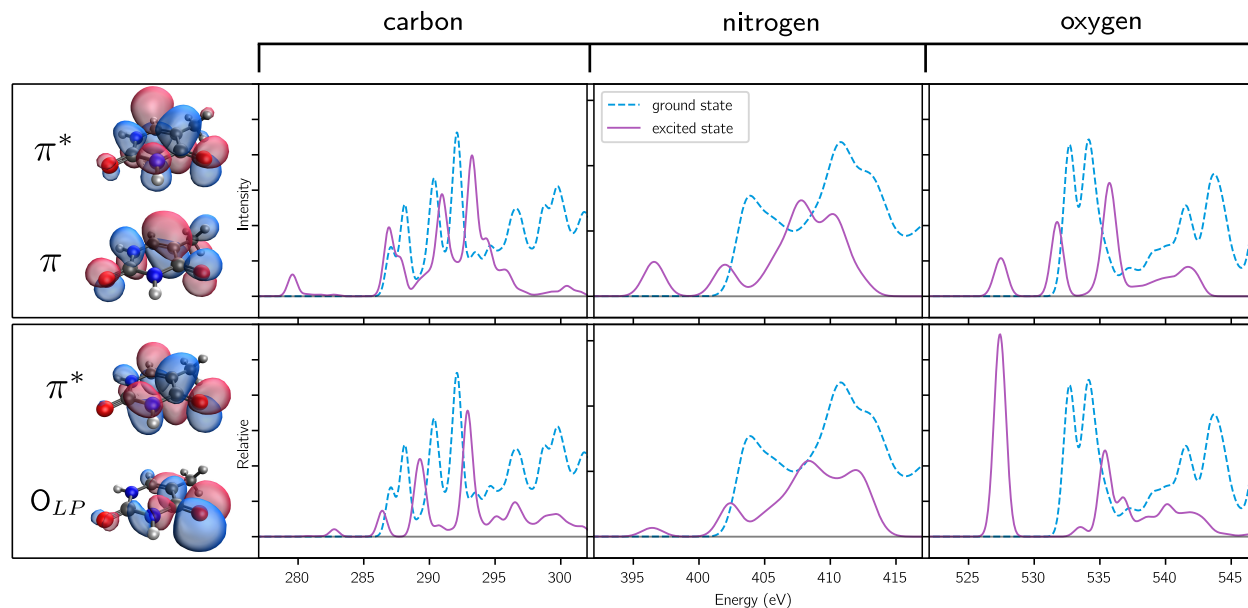
Figure 3.5: 1C-NOCIS NEXAS of the ground state and $O_{LP} \to \pi^*$ and $\pi \to \pi^*$ excited
states of thymine at the C, N, and O K-edges.

correct location in energy and relative intensities of these features remains unknown.

3.4.3   Thymine

Thymine served as a valuable second case study because, aside from being a molecule of
broad chemical and biological interest, it has a variety of heavy atoms, which allowed us
to investigate the behavior of the excited-state NEXAS at different edges. Furthermore,
an experimental TR-NEXAS with sub-100 fs time resolution at the oxygen K-edge was
reported in the literature. [89] Figure 3.5 provides the NEXAS of the $O_{LP} \to \pi^*$ (bottom
panels) and $\pi \to \pi^*$ (top panels) excited states of thymine at the carbon, nitrogen, and
oxygen K-edges, as calculated with 1C-NOCIS 2eOS. With the exception of the oxygen
atom with the dominant contribution to the hole level of the $(O_{LP} \to \pi^*)$ state, and in stark
contrast to water, the calculated ionization thresholds of the two excited states studied here
lie relatively close to those of the ground state for all atoms at the C, N and the O K-edge
(Section S6 of the SI). We speculate that this difference could arise from the fact that the
excited states of thymine feature particle levels with proper valence character, whereas those
of water are all of Rydberg character at the Frank-Condon geometry. [15] Excited states of
Rydberg character are quasi-ionized systems, which are known to feature severe blue-shifts
in their core ionization potentials. [87, 265] In regards to the strongly blue-shifted ionization
potential for one of the oxygens: the $(O_{LP} \to \pi^*)$ excited state migrates electron density
localized at said oxygen atom to a π^* orbital delocalized over the whole molecule. In other

words, it has local charge-transfer character and in a sense this scenario mimicks the effect of a Rydberg excitation or an ionization from the point of view of this oxygen atom.

The c $\to$ SOMO(o) transition at the oxygen K-edge is the most prominent, well-separated feature for identification of the $O_{LP} \to \pi^*$ excited state. This signal is clearly detected in the TR-NEXAS experiment of Wolf *et al.* [89] Direct observation of the bright $\pi \to \pi^*$ state is predicted to be more challenging. One promising way could be to rely on the splitting of the $O_{1s} \to \pi^*$ feature at the oxygen K-edge, which seems to be accentuated in the $\pi \to \pi^*$ state. In theory, then, a positive induced absorption flanking the ground state bleach should be observed with enough time and energy resolution. In the experiment, some induced absorption is observed in this region in the ultra-fast time scales but with the resolution of the TR-NEXAS its not possible to assign it with certainty. The carbon and nitrogen K-edges of the excited states seem to overlap too much with each other and with the ground state at the Frank-Condon geometry to be of use. While thymine is a relatively rigid molecule, it is possible that structural dynamics may result in shifts to the energies and intensities that disentangles the NEXAS of the different states. We relegate an investigation of the spectral consequences of nuclear motion on a more flexible molecule: acetylacetone.

3.4.4 Simulation of the TR-NEXAS of acetylacetone

No set of benchmark excited-state NEXAS exists from theory yet and only a limited number of UV-pump X-ray-probe TR-NEXAS experiments have come out in the last two decades, where the excited state NEXAS are encoded as a differential absorption from the static spectrum. The resolution of this data is compromised, in part, due to the difficulties of constructing instruments capable of the incredible time resolution required to observe these short-lived transient species and, as a result, no high-resolution data coming from experiment exists either. We chose to validate the 1C-NOCIS 2eOS model by generating a theoretical TR-NEXAS for the carbon K-edge of acetylacetone after excitation into S_2 - its lowest $\pi \to \pi^*$ state - for two reasons. First, a TR-NEXAS experiment that reports on these dynamics is present in the literature and it provides among the clearest transients, with a variety of features well separated in energy and time. [90] Second, the excited state dynamics are well-established, providing certainty as to which excited states should be visible in the experiment. [266] Specifically, the $\pi \to \pi^*$ state lives on the order of 50 - 100 fs, decaying into S_1 ($O_{LP} \to \pi^*$). In turn, S_1 lives for hundreds of fs before direct internal conversion into the T_1 ($\pi \to \pi^*$) takes place; the T_2 ($O_{LP} \to \pi^*$) state is found to play a minor role in the relaxation from S_1 into T_1. With $\sim$60 fs X-ray pump pulses, the set up employed by Battacherjee and coworkers to study these dynamics has the time resolution to unambiguously observe the S_1 and T_1 states but not the S_2 state.

Figure 3.6b shows the populations of the S_1 and S_2 states of acetylacetone from a sample AFSSH trajectory, initiated in the bright S_2 state and carried out in the adiabatic basis, with the energies of the states displayed in Figure 3.6a. The system likely remains in the $\pi \to \pi^*$ state in the first 20 fs of the trajectory. Within the subsequent 30 fs, the $O_{LP} \to \pi^*$ state becomes populated. To prepare the construction of a theoretical TR-NEXAS based
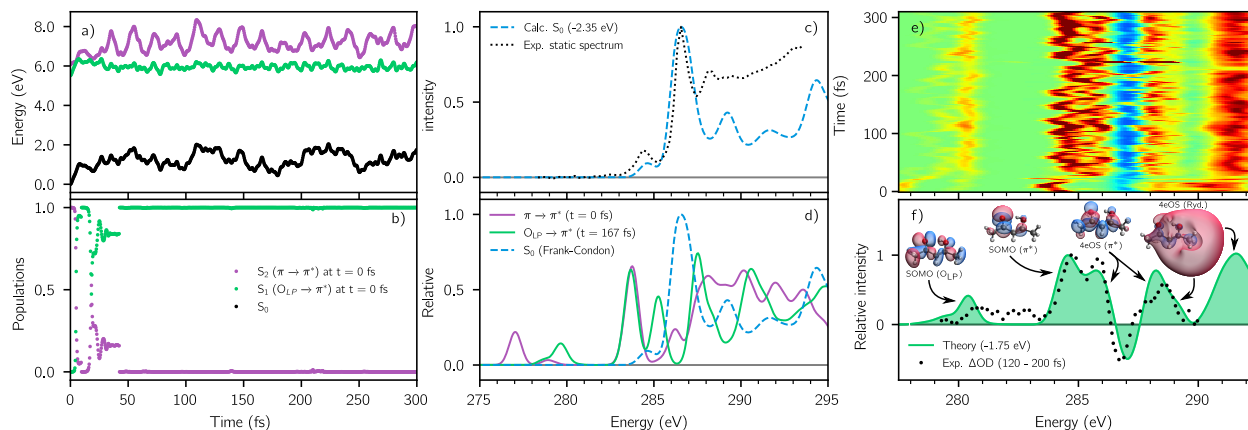
Figure 3.6: a) Energies of the three lowest singlet states (the ground state, the $O_{LP} \to \pi^*$, and $\pi \to \pi^*$ excited states) during an AFSSH NAMD simulation. b) The populations of the two excited states during the NAMD trajectory. c) Comparison of the ground state NEXAS spectrum calculated with 1C-NOCIS, Wigner-broadened with 150 structures, and the experimental NEXAS of Bhattacherjee and coworkers. [90] A shift of -2.35 eV is required to match experiment and theory. d) Comparison of the ground state NEXAS spectrum and the excited-state NEXAS spectra at relevant times, namely when they are populated in the dynamics. e) Simulated TR-NEXAS, shifted by -1.75 eV; see text for details f) Comparison of experimental spectrum, averaged over the 120 - 200 fs time bins compared to the simulated spectrum averaged over the time range while the $O_{LP} \to \pi^*$ is populated in the simulation (50 - 300 fs). [90]

on this trajectory, the ground state spectrum is calculated with standard 1C-NOCIS and compared against the experimental static spectrum, shown in Figure 3.6c. A shift of -2.35 eV, within the errors expected for 1C-NOCIS for closed-shell organic molecules, is required in the computed spectrum to match the dominant feature of the experiment. [114] The shift required represents the remaining dynamic correlation not captured by 1C-NOCIS. Despite a complete disregard to dynamic correlation, this shift is smaller than the +3.50 and -4.25 eV shifts required by DFT-MRCI and RAS-PT2 to align the ground state spectra of butadiene and malonaldehyde, respectively, to either experiment or CVS-ADC(2)-x, and serves as a testament to the dominant relevance of orbital relaxation for a description of core excited states. [251, 252] We proceed by calculating the excited-state NEXAS for S_1 and S_2 with 1C-NOCIS 2eOS out of structures plucked from the NAMD simulations at regular time intervals. An example of the excited state spectra at relevant times is displayed in Figure 3.6d, compared against the calculated ground state spectrum. Finally, the theoretical TR-NEXAS shown in Figure 3.6e is constructed by taking a linear combination of the excited state NEXAS as a function of time, with the coefficients being determined by the NAMD populations and a fraction of the ground-state NEXAS subtracted to simulate the bleach

feature. For the transient spectrum, a smaller shift of -1.75 eV is applied for a better match with the experimental results, obtained from Ref [90].

The impact of the oscillations in the molecular structure during the dynamics on the spectral features are clearly visible in Figure 3.6e. To rigorously account for the dynamics, a larger number of spectral simulations on different trajectories would have to be carried out which would likely blur out the signatures of nuclear motion. While this is beyond the purpose of this part of our work, which is simply to illustrate that full-spectrum calculations based on dynamics can be performed with 1C-NOCIS 2eOS, we carried out an additional spectral simulation that yields similar results (Figure S9).

The effect of the non-adiabatic dynamics on the spectrum within the first 65 fs of the dynamics is displayed in more detail in Figure S8 of the SI. The $C_{1s} \rightarrow SOMO(\pi)$ transitions of the $\pi \rightarrow \pi^*$ state is visible as a mild absorption at 277.5 eV for the first 20 fs of the trajectory. During the same time, the $C_{1s} \rightarrow SOMO(\pi^*)$ transitions manifest as a single intense absorption centered at 284.5 eV. The spectrum changes dramatically at around 25 fs, when a two-peak feature centered at around 285 eV - characteristic of the $O_{LP} \rightarrow \pi^*$ state - emerges and oscillates in and out of the spectra for the remaining 300 fs simulated. Concurrently, the $C_{1s} \rightarrow SOMO(\mathbf{o})$, where $\mathbf{o}$ is now the O_{LP} orbital, blue-shifts by about 2 eV. While the higher-energy regime beyond 288 eV is more crowded for the $\pi \rightarrow \pi^*$ state, there are two well-defined signals for the $O_{LP} \rightarrow \pi^*$ state at 288.5 and 291.5 eV. Averaging the spectral contribution of the $O_{LP} \rightarrow \pi^*$ state results in excellent agreement with the experimental profile (Figure 3.6f)

Figure S9 decomposes in further detail the different contributions to the spectral features of the $O_{LP} \rightarrow \pi^*$ state by visualizing the final particle levels where the probe electrons end up residing as well as the core orbitals from where they emerge. Representative particle levels are visualized in Figure 3.6f. As mentioned earlier, the signals at 278.5 - 280.5 eV arise from the $C_{1s} \rightarrow SOMO(O_{LP})$ transitions out of the keto carbon and the two α carbons with respect to the keto group. They could be the source of the mild positive absorbance in the 281 - 283 eV regime in the experimental spectrum at 120 - 200 fs (which would mean an underestimation by 1C-NOCIS, with the shift that was applied) or else a small fraction of triplet state already present, which absorbs in this regime. Moving up in energy we arrive at the doubled-peaked feature, clearly observed in the experiment. The keto and alcohol carbons give rise to the low-energy peak at 284.3 eV via intense $C_{1s} \rightarrow SOMO(\pi^*)$ transitions. On the other hand, the high-energy peak at 285.8 arises dominantly due to the $C_{1s} \rightarrow 4eOS(\pi^*)$ transition out of the central carbon, where this time the π^* level corresponds to the fully-vacant π^* orbital. The signals beyond arise predominantly from an increasingly dense manifold of Rydberg transitions. The feature from 287.5 to 289.5 eV corresponds in part due to the $C_{1s} \rightarrow 4eOS(\pi^*)$ transition of the alcohol carbon and in part due to $C_{1s} \rightarrow 4eOS(3s, 3p)$ transitions. Finally, the signals beyond 290 eV belong to higher Rydberg states. Since the experimental spectrum reported does not extend to this energy range, it remains for future studies to identify whether this feature is present. Furthermore, this region approaches the onset of the ionization threshold so the agreement with experiment is expected to deteriorate due to the lack of diffuse functions and continuum-treatment techniques.

3.5   Conclusion

To conclude, we highlight the merits of 1C-NOCIS 2eOS by specifying the challenges other
methods would face when attempting to simulate the TR-NEXAS of acetylacetone, as done
in the previous subsection. ΔSCF / TDDFT, the theoretical model employed to calculate
the excited state NEXAS in the study of Bhattacherjee *et al.* requires a shift of 10.3 eV to
match the calculated ground state with experiment and, as explained in Section 3.1.3, the
TDDFT core spectra calculated out of the ΔSCF valence excited states comes with a strong
degree of spin contamination. The clear double-peaked feature present in the experiment
up until 1.5 ps is not observed in any of the computed spectra, likely because ΔSCF /
TDDFT lacks the capacity to appropriately deal with transitions into fully vacant orbitals.
In contrast, 1C-NOCIS requires a shift smaller by a factor of five to match with experiment,
provides excited-state NEXAS free from spin contamination by design, and the resulting
spectral profile matches well with the experimental observation even beyond transitions into
the SOMOs. While hh-TDA DFT may be able to efficiently capture the important c $\rightarrow$
SOMO transitions without spin-contamination, it would require a large shift to align with
experiment and would be incapable to produce the remaining of the spectrum. Most of the
theories able to provide the full spectrum are either too cumbersome (AP-ΔSCF or MS-DFT)
or prohibitively expensive (MR approaches) to simulate a TR-NEXAS for acetylacetone at
the carbon K-edge for a duration on the order of hundreds of femtoseconds. 1C-NOCIS 2eOS,
on the other hand, is an affordable quasi-black-box diagonalization-based approach where
only the valence excited state and the edge of interest need to be specified. Consequently,
it is capable of sampling a range of nuclear configurations efficiently and it is thus amenable
for the TR-NEXAS simulation presented. The only methods capable to produce comparable
results, likely at an increased cost and with perhaps a larger energy shift, would be DFT-
MRCI or RASPT2.

   1C-NOCIS 2eOS is far from a converged theory and there is room for development.
An obvious direction for progress lies in the inclusion of dynamic correlation to further
reduce shifts in the NEXAS to align with experiment and hopefully achieve sub-eV accuracy
as is now plausible for closed-shell systems. Importantly, this could alleviate the possible
differential shifts required for the ground state and the excited states, as was the case for
our TR-NEXAS simulation of acetylacetone. A generalization of 1C-NOCIS 2eOS to DFT,
as electron-affinity (EA) TDDFT did for EA-CIS for closed-shell systems, is an attractive
candidate. [114, 267] Alternative options within the DFT framework could be taking lessons
from (s)-TDDFT or simply employing DFT orbitals in the 1C-NOCIS 2eOS model *a la* DFT-
MRCI. [230, 241–243, 250] Alternatively, 1C-NOCIS 2eOS could be corrected for dynamic
correlation within a wave function framework. The recently-studied seniority- and hierarchy-
CI sets an excellent framework for exploration, and advances in this direction will require
maintaining a balanced treatment of the correlation in the initial valence excited state as well
the core excited states. [257, 258] Furthermore, since 1C-NOCIS 2eOS would ideally rely on
NAMD simulations to simulate the excited-state NEXAS, a generalization to DFT would
allow for the use of TDDFT to calculate the valence excited states themselves, and lead to

NAMD trajectories and structures of a quality better than CIS. An alternative would be to
simply use TDDFT NAMD as a source for the structures and populations but a mismatch
in the level of theory employed for the dynamics and for the generation of the excited state
NEXAS presents a book-keeping challenge, since the ordering of the valence excited states
may be different in TDDFT and in CIS. Another exciting avenue for development is the
inclusion of spin-orbit coupling into the theory to calculate NEXAS beyond the K-edge.
This would allow 1C-NOCIS 2eOS to serve TR-NEXAS experiments like those carried out
recently at the iodine $N_{4,5}$-edge. [268–270]

Chapter 4

Femtosecond core-level spectroscopy reveals signatures of transient excited states in the photodissociation of Fe(CO)$_5$

4.1 Introduction

The promise of harnessing sunlight to drive thermodynamically unfavorable chemical reactions with high degree of selectivity has made photocatalysis an active and growing area of research. Organometallic photocatalysts are central to these investigations, where electronic excitations of ligand-to-metal (LMCT) or metal-to-ligand charge transfer (MLCT) character induce changes in the oxidation or coordination state of the transition metal center that can subsequently catalyze a range of chemical reactions. [271–281] A bottom-up assessment of organometallic systems for photocatalysis applications first requires understanding their inherent excited state dynamics in isolation, as photocatalytic activity strongly depends on electronic properties such as excited state lifetimes and intersystem crossing (ISC) rates. Of the numerous organometallic photocatalysts, transition metal carbonyls are textbook examples for understanding metal-ligand bonding and excited state photochemistry. [282–284] In these systems, excitation of the MLCT transition leads to the cleavage of metal-CO bonds to form coordinatively unsaturated and catalytically active transition metal centers.

We report on the ultrafast photochemistry of gas-phase iron pentacarbonyl, Fe(CO)$_5$, a model photocatalyst that has garnered much interest over the last few decades as a useful proxy for studying photocatalytic activity and metal-ligand bonding in excited states. [285–297] It is now well established that 266-nm excitation of Fe(CO)$_5$ leads to the formation of iron tricarbonyl, Fe(CO)$_3$, through an iron tetracarbonyl intermediate, Fe(CO)$_4$. The latter was recently detected via ultra fast time-resolved infrared (TR-IR) spectroscopy. [298] However, open questions remain about the dissociation dynamics due to the high density of

electronic states and the ultrafast timescales on which these processes take place. [299–305]
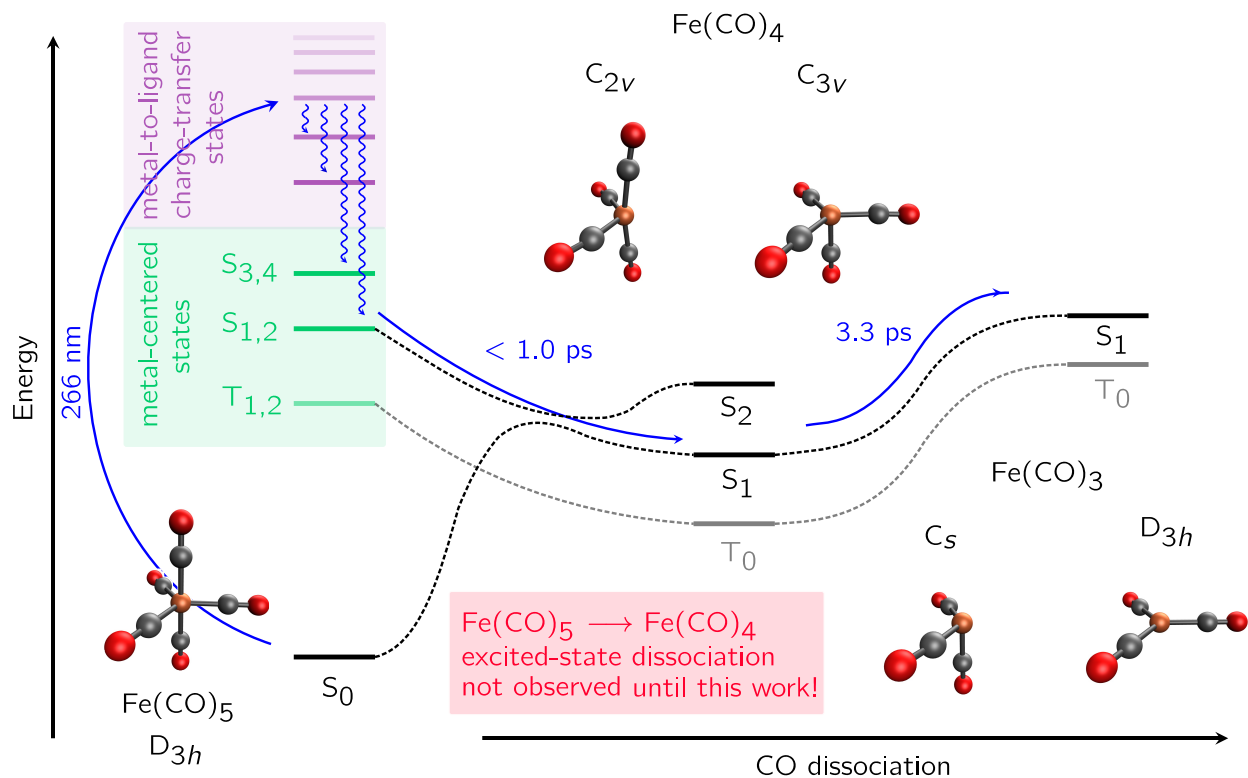
Figure 4.1: Photodissociation pathway of $Fe(CO)_5$: what is established in the literature and what remains to be observed.

The current picture for the excited state pathway underlying $Fe(CO)_5$ photodissociation in the gas phase, illustrated in Figure 4.1, originates from a significant body of work. [298, 299, 305–310] A 266-nm photon is believed to excite $Fe(CO)_5$ into its three lowest bright MLCT states followed by oscillatory internal conversion (IC) into four metal-centered (MC) states. [305, 307] Dissociation, predominantly of the axial ligand, is theorized to occur on the MC surfaces, accompanied by structural distortion along scissoring modes between the ligands. [305] When on the lowest dissociative MC state, the intersection seam with the $^1A_1$ closed-shell singlet state is estimated to become accessible within 2.7 - 3.0 Å of Fe-CO elongation. This intersection is guaranteed by symmetry at a T_d geometry for the fully dissociated product $Fe(CO)_4$. [298, 307] Note that after passage through the intersection seam, literature has assumed the closed-shell surface to remain the lowest singlet surface in $Fe(CO)_4$, but a rigorous assessment of the character of this state has not been made. All in all, the excited state dynamics were suggested to occur with sub-100 fs time constants.

[307] In part because of the ultra-fast timescales of these processes, a direct spectroscopic observation of the excited state dissociation remains absent.

Fe(CO)$_3$ is subsequently formed on a 3-ps timescale from dissociation of Fe(CO)$_4$. [298, 309, 310] Owing to the triplet ground electronic states in Fe(CO)$_4$ and Fe(CO)$_3$, Trushin *et al.* had suggested the possibility of nanosecond ISC as an alternative pathway for dissociation. [307] However, picosecond electron diffraction and core- and valence-photoelectron spectroscopy eliminated the role of the triplet manifold in the gas-phase photodissociation dynamics. [308, 310] There is not enough internal energy in the system afterwards for further loss of an additional ligand in the singlet manifold. [298]

We present ultrafast extreme ultraviolet transient absorption spectroscopy (XUV-TRAS) of gas-phase Fe(CO)$_5$ near the Fe M$_{2,3}$-edge after excitation at 266 nm, revealing core-to-valence transitions in the 52-67 eV range evolving with 100-fs and 3-ps time constants. We calculate the core-level absorption spectra of the relevant valence excited states using the one-centered non-orthogonal configuration interaction singles (1C-NOCIS) theory, recently adapted to handle two-electron open-shell (2eOS) singlet references, [184] and compare them against the experimental XUV-TRAS using a simplified model of the excited state dynamics during photodissociation. To identify the spectroscopic signatures of Fe(CO)$_5$, Fe(CO)$_4$, and Fe(CO)$_3$, we use electron-affinity time-dependent density functional theory (EA-TDDFT), which uncovers the dramatic role of orbital relaxation in Fe M$_{2,3}$-edge spectra of these species. [114, 267] This closely integrated experiment-theory work thus reports the first spectroscopic signatures of valence-excited state dynamics of Fe(CO)$_5$ to form Fe(CO)$_4$ and Fe(CO)$_3$ unveiling the role of coupled nuclear and electronic motion along the dissociation pathways.

4.2 Results and Discussion

4.2.1 Excited states of Fe(CO)$_5$ at the Frank-Condon geometry

The four singlet excited states of Fe(CO)$_5$ responsible for dissociation correspond to metal-centered (MC) excited states, where an electron from any of the four occupied $3d$ orbitals on the metal (d$_{xz}$, d$_{yz}$, d$_{xy}$, or d$_{x^2-y^2}$) is promoted to the only empty $3d$ orbital (d$_{z^2}$). Table 4.1 shows the rCAM-B3LYP/aug-pcseg-2 TDDFT excited states within the TDA approximantion, including scalar relativistic effects via the X2C model. In the 3.9 - 4.4 eV region there is already 5 states. In other words, the manifold of states is dense and establishing the exact ordering at the Frank-Condon geometry is beyond the accuracy of TDDFT.

With this limitations in mind, our model tells us that the two lowest singlets correspond to a degenerate pair of $^1E'$ states with d$_{x^2-y^2} \rightarrow$ d$_{z^2}$ and d$_{xy} \rightarrow$ d$_{z^2}$ character. As the change in Mulliken charge on the iron atom (Fe Δq) suggest, little charge flows out of the iron center during this excitations which reflects their nature as metal-centered states. S$_{3-6}$ correspond to different dark d$\rightarrow \pi^*_{CO}$ metal-to-ligand charge transfer (MLCT) states with larger Fe Δq. S$_7$ and S$_8$ correspond to the second pair of degenerate MC states, of d$_{xz} \rightarrow$ d$_{z^2}$ and d$_{yz} \rightarrow$

d$_{z^2}$ character. S$_9$ and the degenerate pair at S$_{10}$ and S$_{11}$ corresponds to the first three bright
MLCT states, which are likely being activated when the system is irradiated with light at
266 nm.

Table 4.1: First 10 singlet excited states of Fe(CO)$_5$ (D$_{3h}$).

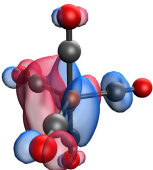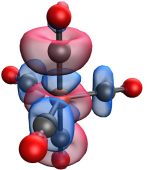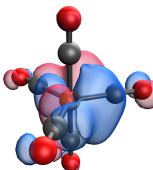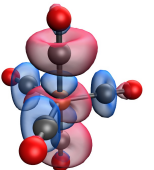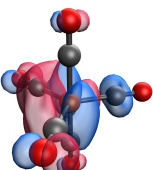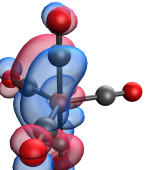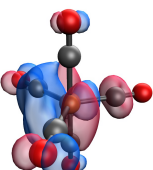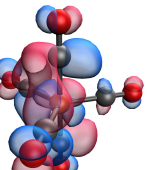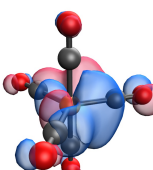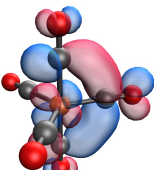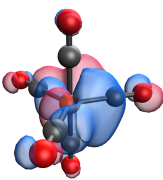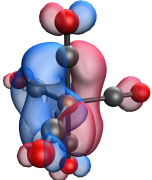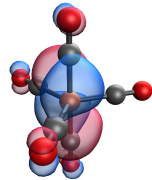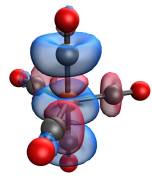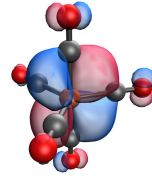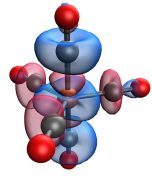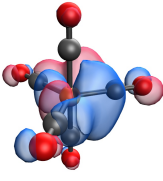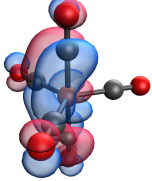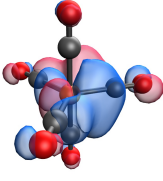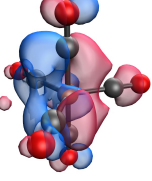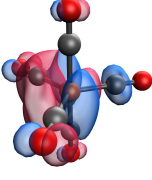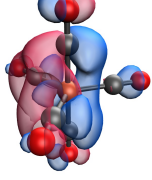| State	Symm.	Energy	O.S.	Fe Δq	Character	HONTO	LUNTO
S$_1$	$^1E'$	3.956	0.000	0.080	MC		
S$_2$	$^1E'$	3.956	0.000	0.080	MC		
S$_3$	$^1A_1''$	4.096	0.000	0.514	MLCT		
S$_4$	$^1E''$	4.382	0.000	0.467	MLCT		
S$_5$	$^1E''$	4.382	0.000	0.467	MLCT		

Table 4.2: Continuation of Table 4.1

State	Symm.	Energy	O.S.	Fe Δq	Character	HONTO	LUNTO
S_6	$^1A_1'$	4.752	0.000	0.505	MLCT		
S_7	$^1E''$	4.753	0.000	0.175	MC		
S_8	$^1E''$	4.753	0.000	0.176	MC		
S_9	$^1A_2''$	4.772	0.083	0.512	MLCT		
S_{10}	$^1E'$	4.910	0.049	0.504	MLCT		
S_{11}	$^1E'$	4.910	0.049	0.504	MLCT		

As will be elaborated on in Subsection 4.2.2, the four MC states are dissociative with respect to the loss of a ligand, so they will quickly become the four lowest in energy upon departure of the Frank-Condon region. For convenience, then, we will refer to the MC states according to their energy ordering on dissociation.

- S_1: $d_{x^2-y^2} \rightarrow d_{z^2}$

- S_2: $d_{xy} \rightarrow d_{z^2}$

- S_3: $d_{xz} \rightarrow d_{z^2}$

- S_4: $d_{yz} \rightarrow d_{z^2}$

4.2.2 Computational study of the different iron carbonyls and the structures connecting them

Fe(CO)$_5$: ultra-fast non-adiabatic dynamics near the FC region

To prepare for the construction of a theoretical time-resolved XAS spectrum to be compared with the experimental results we make use of the non-adiabatic molecular dynamics (NAMD) simulations carried out by Banerjee and coworkers of the relaxation of Fe(CO)$_5$ after excitation into its lowest bright MLCT state. [305] I gratefully acknowledge them for providing me with their NAMD trajectories (structures, populations) for this study. Their dynamics (presented in Figure 4.2) reveal that, after initalization in the first bright MLCT state (S_7 in their calculations), a majority of the population has decayed into the MC states (S_{2-5} in their calculations) by the first 200 fs. By 400 fs, virtually all of the surviving trajectories find themselves in the lowest MC state.

Naturally, dissociation of the first CO ligand on the excited state surface takes place as the MC states become populated. Since their dynamics are carried out with TDDFT and the closed-shell singlet surface is not included in the simulations due to a failure of the method for conical intersections with closed-shell surfaces, [311] the dynamics involving the latter remain to be fully established. However, based on the proximity between the lowest MC and the closed-shell singlet surfaces (< 0.1 eV) once the Fe$-$CO distance goes beyond 2.7 - 3.0 Å, I estimate non-adiabatic transitions into the closed-shell singlet to become accessible at this point.

You may notice that the lowest populated state in Figure 4.2 is that corresponding to S_2 at the Frank-Condon geometry. This arises from the fact that there is a dark MLCT state (S_3 in Table 4.1) which ends up being the lowest in energy at the Frank-Condon geometry in their calculations. As elaborated on Subsection 4.2.1 the density of states is high and the exact energetic ordering is uncertain. Regardless, the MC states will become the lowest in energy shortly after dissociation begins since the MLCT states are all bound. As a result, whether the exact energy location of the low-lying MLCT state is below the first MC states or not may be relatively inconsequential to the photodissociation dynamics.
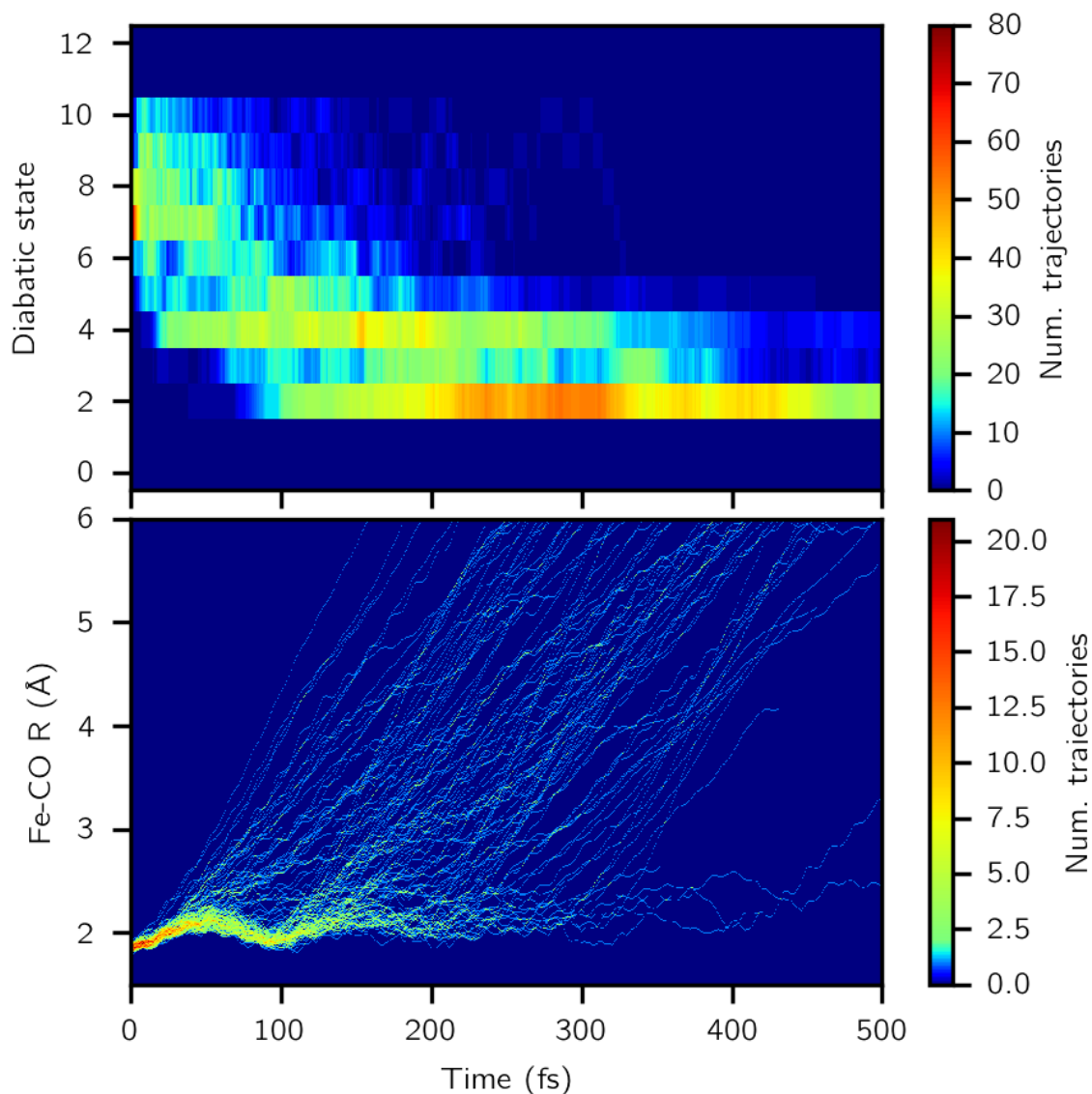
Figure 4.2: Non-adiabatic molecular dynamics of Banerjee *et al.* [305]

Fe(CO)$_5$: exploring key structural degrees of freedom

A 2D potential energy surface scan along the axial Fe–CO bond lengths and the angle
formed by the equatorial ligands and the Fe center, while relaxing the remaining internal
degrees of freedom to their minima in S$_0$, reveals two key features of the lowest MC excited
states. As Figure 4.3 makes clear, the MC states are dissociative with respect to the axial
ligand. Furthermore, they are Jahn-Teller unstable with respect to motion along modes of e$'$
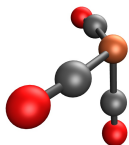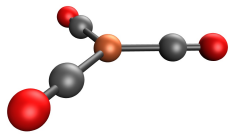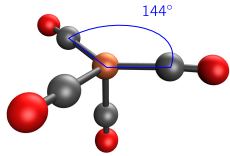
Figure 4.3: Potential energy surfaces for $S_{0\text{-}2}$ of Fe(CO)$_5$ along the Fe$-$CO$_{ax.}$ bond length and the CO$_{eq.}-$Fe$-$CO$_{eq.}$ angle.

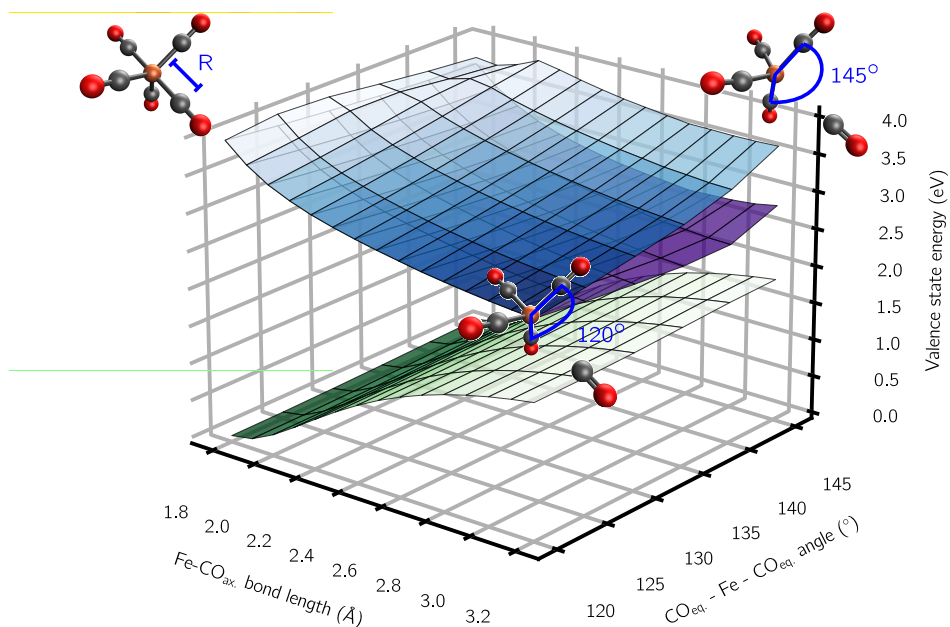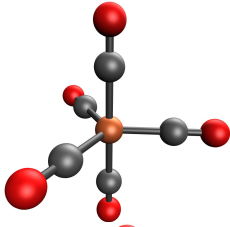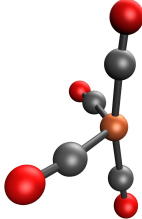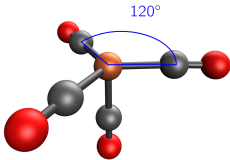symmetry. For example, scanning the CO$_{eq.}-$Fe$-$CO$_{eq.}$ scissoring mode lifts the degeneracy between the two states. The same phenomena holds for the other set of MC states (S$_{3, 4}$) because they are also of ^1E symmetry within C$_{3v}$. As a result, an Fe(CO)$_5$ molecule losing an axial CO ligand while traversing on the MC states will distort away C$_{3v}$ symmetry as the JT pairs of ^1E MC states break into $^1B_2$ and $^1A_2$ states within C$_{2v}$.

Fe(CO)$_4$ and Fe(CO)$_3$: key structures on the closed-shell singlet surface

Table 4.3 summarizes the key structural configurations during and after the loss of CO ligands, namely the minimum energy crossing point (MECP) that connects the two lowest singlet surfaces and the structural minima after dissociation. As corroborated by an explicit MECP optimization using spin-adapted spin-flip (SA-SF)-TDDFT, [312] the two surfaces for the Fe(CO)$_4$ photoproduct touch at their MECP in a structure of C$_{2v}$ symmetry (last row of Table 4.3). Importantly, this relaxes the high-symmetry requirement of IC into the lowest singlet surface at a T$_d$ geometry proposed in the literature. [307] Furthermore, as visualized in Figure 4.1, the lowest triplet surface becomes the formal ground state of Fe(CO)$_4$ either during or after dissociation. Therefore, the ground singlet state of Fe(CO)$_5$ correlates with what is formally an excited singlet state in Fe(CO)$_4$ and whether the lowest singlet remains

Table 4.3: Top: relative energies for the minima in the ^{1}A closed-shell singlet surface. Bottom: structure for the MECP between the $^1A_1$ and $^1B_2$ states of Fe(CO)$_4$.

Species	Structure	rCAM-B3LYP	ωB97M-V	CCSD(T)
Fe(CO)$_5$ (D$_{3h}$)		0.000	0.000	0.000
Fe(CO)$_4$ (C$_{2v}$)		1.799	1.800	1.899
Fe(CO)$_4$ (C$_{3v}$)		1.852	1.804	1.885
Fe(CO)$_3$ (C$_s$)		3.535	3.437	-
Fe(CO)$_3$ (D$_{3h}$)		3.877	3.780	4.184
Fe(CO)$_4$ (C$_{2v}$)		$\leftarrow$ $^1A_1$ / $^1B_2$ MECP structure.		

the closed-shell $^1A_1$ state or the open-shell $^1B_2$ state remained an open question.

$$T_1^{Fe(CO)_5} \xrightarrow{\text{ligand dissociation}} T_0^{Fe(CO)_4} \tag{4.1}$$

$$S_0^{Fe(CO)_5} \xrightarrow{\text{ligand dissociation}} S_1^{Fe(CO)_4} \quad \text{(closed-shell or open-shell?)} \tag{4.2}$$
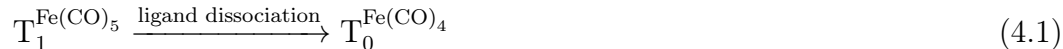
Starting from the Frank-Condon geometry, I was only able to find what turned out to be non-observable minima in the lowest open-shell singlet surface of $Fe(CO)_4$ and $Fe(CO)_3$. The key signature that classified these structures as "fantastical" were anomalously large second-derivatives at the presumed minima after a successful geometry optimization (vibrational frequencies on the order of 35,000 cm−1!). This strong curvatures relates to a breakdown of the Born-Oppenheimer approximation and suggests that the structure is within a region of strong non-adiabatic coupling. This led to a wonderful paper I was a part of but is not in this thesis. [186]

The only genuine minima I could find were on the closed-shell surface and they will be discussed shortly. Before, let me mention an analysis of the electronic wave function that we carried out at these minima with the complete active space self-consistent field (CASSCF) method to identify the degree of strong correlation present in the unsaturated iron carbonyls. The adaptive sampling configuration interaction (ASCI) solver was used for CASSCF to ensure the tractability of a comprehensive active space. Specifically, we incorporated all the valence orbitals in the iron center (3d, 4s, 4p, 4d) as well as those in the CO ligands (σ, π, π^*, and σ^*) to ensure a reliable assessment. Furthermore, we followed the protocol detailed in Ref. [313] to avoid the local minima problem inherent to selective CI approaches. The active space is summarized in Table 4.4

Table 4.4: CASSCF active spaces

Fragment	e^-	orb.
Fe	$3d(6) + 4s(2) = \mathbf{8}$	$3d(5) + 4s(1) + 4p(3) + 4d(5) = \mathbf{14}$
$(CO)_x$	$\sigma(2) + \pi(4) = \mathbf{6x}$	$\sigma(1) + \pi(2) + \pi^*(2) + \sigma^*(1) = \mathbf{6x}$

According to the contributions of the closed-shell configuration C_0 - a measure of strong correlation - of the different unsaturated iron carbonyls, we conclude them all to be of single-reference character at their equilibrium geometries (Table 4.5). [314] As a result, we proceed with confidence with our use of single-reference methods for the purposes of reporting energies and generating theoretical spectra. Let's now move to a discussion of the structures themselves.

For $Fe(CO)_4$, the geometry optimizations uncover a structure of C_{3v} symmetry in addition to the well-characterized C_{2v} structure. [299, 300, 308] Our calculations predict the C_{3v} structure for $Fe(CO)_4$, which has only been briefly alluded to by Daniel *et al.*, [299] to be nearly isoenergetic with the C_{2v} structure, with only a small barrier of less than 0.1 eV

Table 4.5: Contributions of the dominant configuration to the normalized CASSCF / aug-pcseg-1 wave functions for the systems under study.

Structure	C_0
$Fe(CO)_4$ (C_{3v})	0.864
$Fe(CO)_4$ (C_{2v})	0.867
$Fe(CO)_3$ (D_{3h})	0.867
$Fe(CO)_3$ (C_s)	0.862

for interconversion. Due to the large amount of internal energy remaining in the $Fe(CO)_4$ fragment after first CO loss, $Fe(CO)_4$ likely exists in strongly distorted geometries oscillating between the C_{2v} and C_{3v} minima. A question that remains unclear to me is why this structure was not identified in the seminal ultra fast electron diffraction (UED) experiments of Ihee and co-workers, which only seem to reveal a structure of C_{2v} symmetry. [308] One possibility could be that their experiment, which uses two-photon 620 nm (i.e. 310 nm) excitation and records the UED 250 ps after excitation, is actually observing the C_s structure of $Fe(CO)_3$ which they did not consider when analyzing the spectra.

Summary of predicted pathway

A summary of the predicted photodissociation pathway in the literature, [305, 307] enhanced in this study, is as follows. Irradiation of $Fe(CO)_5$ with light at 266 nm causes photoexcitation into the lowest bright MLCT states. Subsequently, ultra-fast IC conversion (on the order of 100 fs) is predicted to occur into the MC states. The surfaces defined by these states cause $Fe(CO)_5$ to lose an axial ligand and distort along scissoring bending modes involving the remaining ligands. Internal conversion into the lowest singlet surface, which is of closed-shell character but is at this point formally an excited state as a result of the triplet ground state, occurs when a structure finds itself in the lowest MC surface and the dissociating ligand is about 2.70 - 3.0 Å away. When on the closed-shell singlet surface, $Fe(CO)_4$ exists as strongly dissorted structures with a large amount of internal energy, oscillating about the two minima of C_{2v} and C_{3v} symmetry. Finally, the internal energy causes $Fe(CO)_4$ to lose an additional ligand, yielding $Fe(CO)_3$. The lowest minima of the latter is of C_s symmetry, with another minima of D_{3h} symmetry higher in energy.

4.2.3 Simulations of the XUV time-resolved absorption spectra

With a detailed understanding of the dynamics after photoexcitation at hand, we proceed to predict the XUV absorption signatures of the relevant species to compare with the experimental result. In the experiment, the static XUV absorption spectrum of $Fe(CO)_5$ in the gas

phase serves as a reference against which the changes in optical density (ΔOD) are recorded in a time-resolved manner after photoexcitation with 266 nm UV light. Correspondingly, the calculated ground-state XUV spectrum of Fe(CO)$_5$ serves as a computational reference to compare with the calculated XUV spectrum of the excited states of Fe(CO)$_5$, as well as Fe(CO)$_4$ and Fe(CO)$_3$ in their closed-shell singlet minima.

XUV absorption spectrum of ground state Fe(CO)$_5$
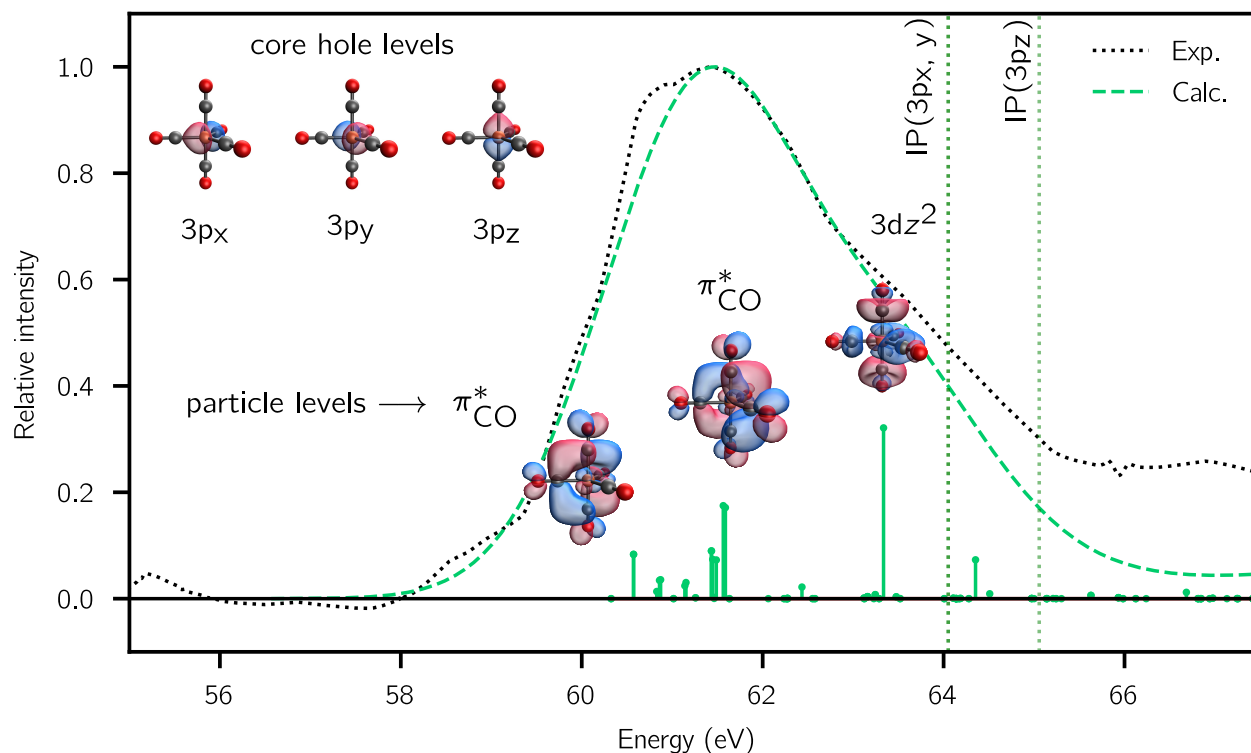
Figure 4.4: Absorption spectrum of ground-state Fe(CO)$_5$.

Figure 4.4 displays the XUV absorption spectrum of ground-state Fe(CO)$_5$ near the Fe M$_{2,3}$-edge, abbreviated as M-edge from now on. The experimental spectrum shows a sharp absorption feature centered at 61 eV with a full-width-at-half-maximum (FWHM) of 4 eV and a shoulder stretching to 70 eV. The theoretical XUV absorption spectrum, calculated with EA-TDDFT using the rCAM-B3LYP functional, is in good quantitative agreement with experiment and predicts the feature near 61 eV to arise predominantly from excitations of $3p \rightarrow \pi^*_{CO}$ character, and the shoulder at 64 eV to arise from a strong $3p_z \rightarrow d_{z^2}$ transition. The features above 65 eV are attributed to transitions into continuum states above the ionization energy.

Excited state XUV absorption spectra during dissociation
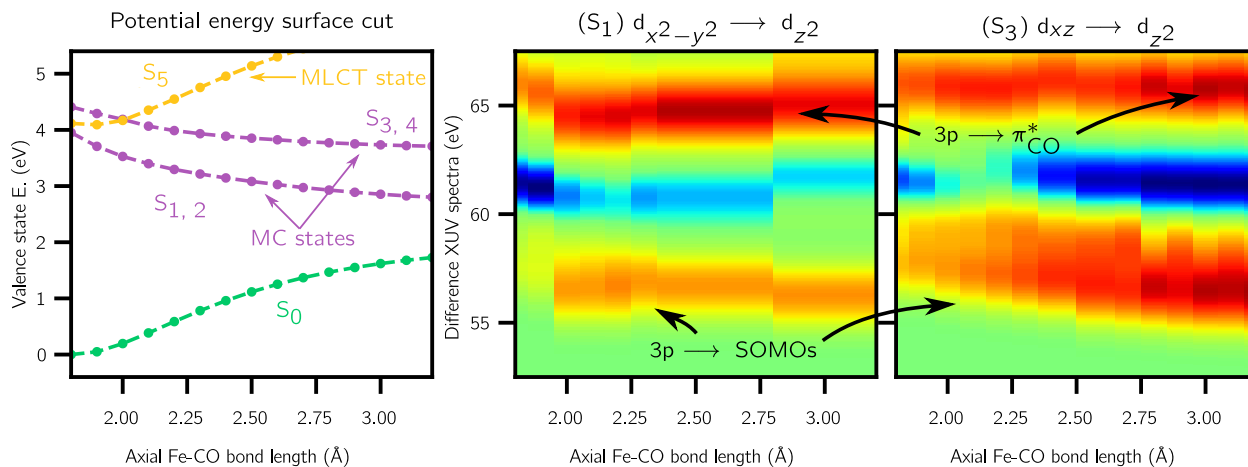
Figure 4.5: Valence excited-states (left) and XUV absorption of the S_1 and S_3 valence excited states (center and right) relative to the ground state as a function of axial Fe$-$CO distance.

Taking a cut along perfect axial dissociation without distorting the CO$_{eq.}$$-Fe-CO_{eq.}$ angle provides the simplest possible dissociation pathway. The energies of the lowest singlet states for this pathway are shown in the left-most panel of Figure 4.5. From these structures, we use 1C-NOCIS 2eOS (Chapter 3) to calculate the M-edge absorption spectra of the MC excited states (S$_{1\text{-}4}$) and find that their spectra show striking differences with respect to ground-state Fe(CO)$_5$. To illustrate, the difference spectra for one of each of the two pairs of MC states with respect to Fe(CO)$_5$ in its ground state are shown in the right two panels of Fig. 4.5. The transitions into the singly-occupied-molecular orbitals (SOMOs) - signatures of electronic states with unpaired electrons - appear as a broad induced absorption red-shifted by roughly 4-5 eV relative to the dominant ground-state Fe(CO)$_5$ feature. Furthermore, the $3p \to \pi^*_{CO}$ transitions, analogous to those observed for ground-state Fe(CO)$_5$, shift by 5 eV to higher energies. The $3p \to \pi^*_{CO}$ transitions in the 63-70 eV range show subtle differences in center photon energies for different MC states: while the $3p \to \pi^*_{CO}$ transitions in S$_{1,2}$ are centered at 64.5 and 65 eV, they are centered at 65.9 eV in S$_{3,4}$. Select outlines are provided in Figure S14 of the paper.

Simulating a XUV-TRAS for Fe(CO)$_5$: challenges

The ideal way to simulate the transient spectrum would involve the actual structures from the NAMD and include contributions due to MLCT states as well as the MC states, similar to the exercise carried out for for acetylacetone in Section 3.4.4. Two challenges that prevented us from doing so are explained below.

The MC states can be described by a single NTO pair (Section 1.3.2), where the lowest-unoccupied(LU)NTO are associated with either the $d_{x^2-y^2}$, d_{xy}, d_{xz}, or d_{yz} hole states and the highest-occupied (HO)NTO is associated with the d_{z^2} particle state. As a result, they can be described by ROKS (Section 1.3.5), upon which 1C-NOCIS 2eOS relies for a description of the valence excited state. On the other hand, the MLCT states exhibit genuine configuration interaction and condense into *two* NTO pairs after a singular value decomposition of the excited-state CIS wave function, thus placing them beyond the ability of 2eOS ROKS to capture. For this reason, we can only account for the contribution to the transient spectra due to the MC states.

The inclusion of the MLCT states aside, the correct approach to simulate the transient would have been to carry out NAMD with CIS or ROKS(HF) given that these are the valence excited states employed in the 1C-NOCIS 2eOS calculations for the XUV spectrum. However, Fe(CO)$_5$ is too complex a system and correlation may be essential, so that dynamics from an uncorrelated theory may be qualitatively incorrect. Directly using the structures from TDDFT NAMD simulations is not possible because, given the high density of states, there is not a straight-forward way to map the adiabatic states of TDDFT and those of CIS or ROKS(HF) at a fixed geometry.

Simulating a XUV-TRAS for Fe(CO)$_5$: approach taken

An approximate TR-XUV spectrum was constructed instead by relying on the calculated XUV spectra for the four MC states as a function of bond distance (Figure 4.3) and the NAMD dynamics as follows. First, we establish the relationship between the diabatic character of the four MC states in the simulations of Banerjee *et al.* (Table 4.6).

Table 4.6: Mapping between diabatic state and state character in the FC geometry.

diabatic state in NAMD	S$_2$	S$_3$	S$_4$	S$_5$
character at FC geometry	$d_{x^2-y^2} \to d_{z^2}$	$d_{xy} \to d_{z^2}$	$d_{xz} \to d_{z^2}$	$d_{yz} \to d_{z^2}$
symmetry (C$_{2v}$)	$^1B_2$	$^1A_2$	$^1B_2$	$^1A_2$

Second, we characterize each time step in a NAMD trajectory by a) the maximum Fe-CO axial bond-length and b) the diabatic state (S$_1$, S$_2$, S$_3$, S$_4$) populated, as previously shown in Figure 4.2. To each time step, we assign a XUV difference spectrum as that corresponding to the surrogate, idealized structure from our approximate dissociation model (Section 4.2.3) in the same diabatic state. Finally, the spectral contribution from each trajectory is averaged. Crucially, we switch to the 1C-NOCIS difference spectra of Fe(CO)$_4$ in its closed-shell singlet state $^1A_1$ for a trajectory in the lowest MC state ($^1B_2$) that has fully dissociated (i.e. Fe–CO bond length > 2.7 Å). We use an equal-weight superposition of the calculated XUV absorption spectra of Fe(CO)$_4$ in both C$_{2v}$ and C$_{3v}$ geometries and empirically scale the resulting difference spectra with ground-state Fe(CO)$_5$ by 6 to match the experimental behavior of the ground state bleach (no change beyond the IRF of the

instrument). The simulated XUV time-resolved absorption spectra (TRAS) is compared against experiment in the followin Subsection.

4.2.4 Comparison of the simulated XUV TRAS with experiment

Ultra fast timescales: signatures of the MC excited states

Figure 4.6: Experimental (left) and theoretical (center) XUV-TRAS for Fe(CO)$_5$ photodissociation. The time evolution of the 65 eV feature is presented on the right panel.

These experiments were carried out by Jan Troß and Krupa Ramashesa at Sandia National Laboratory, and they generously allowed me to include the results in this thesis for the purposes of comparing with my simulations. In the experiment, shown in the left-most panel of Figure 4.6, photoexcitation at 266 nm causes an abrupt change in the XUV absorption spectrum. The main bleach, or decrease in the optical density (ΔOD), at 61.5 eV is associated with the loss of ground-state Fe(CO)$_5$ and additional weak bleach features are present at 50 and 70 eV. Induced absorption (increase in ΔOD) features are centered at photon energies of 54, 58.5 and 65 eV, each with distinct temporal behavior. The temporal lineouts are presented in Section 1.3 in the SI of the paper. The bleach at 61.5 eV appears promptly and reaches its asymptote within the experimental instrument response function (IRF) of 74 fs FWHM. The induced absorption at 65 eV shows a short-lived 1 mOD signal that decays to an asymptote of 0.5 mOD on a 100-fs timescale, as observed in the right-most panel of Fig. 4.6. Furthermore, within the same timescale, it narrows in full-width from about 5 to 1 eV, mostly losing intensity on the higher energy side of the feature. The induced absorption at 58.5 eV reaches 90% of its asymptotic value by 400 fs, whereas the feature centered at 54 eV significantly grows on a picosecond timescale; by 400 fs, this feature has only reached 50% of its asymptotic value.

The center panel of Figure 4.6 shows the simulated XUV-TRAS derived from our dissociation model and excited state spectral calculations (Fig. 4.3), convolved along the time axis with a Gaussian of $\sigma = 44$ fs to simulate the instrument response function. The simulated XUV-TRAS describes key features of the first 400 fs of our experimental XUV-TRAS: the ground state bleach at 61.5 eV, the intense induced absorption at 58.5 eV, and the weak, short-lived induced absorption centered at 65 eV. The lack of induced absorption below 55 eV in the theoretical XUV-TRAS may be a consequence of our simplified dissociation model, which ignores the structural deformation caused by the excited-state dynamics. Importantly, the simulated XUV-TRAS show the narrowing of the 65 eV feature on the timescale of 100 fs, as observed in the experiment. The evolution of this signal arises from the decay of the population in the MC states into the closed-shell singlet during dissociation of the first CO ligand. The reason for this decays is the strong absorption of the MC states at 65 eV, arising from intense $3p \rightarrow \pi_{CO}^*$ transitions. On the other hand, as will be elaborated on the following Subsection, the unsaturated iron carbonyls on the closed-shell singlet surface do not feature strong absorption at 65 eV. This major finding constitutes the first spectroscopic observation of the electronic excited state dynamics during the photodissociation of Fe(CO)$_5$.

Longer timescales: signatures of Fe(CO)$_4$ and Fe(CO)$_3$

Given the rapid IC onto the closed-shell singlet surface along the first CO loss, now S$_1$ due to the triplet ground state, we proceed by considering only the spectral contributions of the unsaturated iron carbonyls Fe(CO)$_4$ and Fe(CO)$_3$ on S$_1$ for comparison with experiment. The left panel of Figure 4.7 shows the XUV absorption spectra of the unsaturated iron carbonyls, calculated with EA-TDDFT at the indicated geometries. While these are idealized structures, since in reality the leftover energy delivered by photoexitation remains within the photofragments as internal energy, a clear trend emerges. Fe(CO)$_4$ is predicted to absorb up to 5 eV lower in energy than Fe(CO)$_5$. Fe(CO)$_3$ is predicted to also absorb in this regime but with higher intensity.

To enable qualitative assignment of experimental spectra using EA-TDDFT spectra of the unsaturated iron carbonyls, my experimental collaborators performed a global target analysis of the experimental XUV-TRAS to extract spectral components and time constants. Since sequential dissociation

$$\text{Fe(CO)}_5 \xrightarrow{\text{excited-state dissociation}} \text{Fe(CO)}_4 \xrightarrow{\text{thermal dissociation}} \text{Fe(CO)}_3 \qquad (4.3)$$
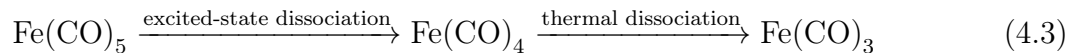
is the only active mechanism in the 266 nm photodissociation of gas-phase Fe(CO)$_5$, they employed a sequential global kinetic model to fit the experimental XUV-TRAS spectra. The fit, which incorporates the instrument response function, produces a sequence of exponential decays with time constants of 120±2 fs, 3300 fs, and 1 μs time constants. The 3300 fs and 1 μs time constants were fixed to align with previous work and to account for the signal that persists within the experimental time window. [298, 307, 310]) The global target analysis generates evolution-associated difference spectra (EADS), which isolate spectral changes corresponding to each time constant, allowing assignment of these spectra to contributions
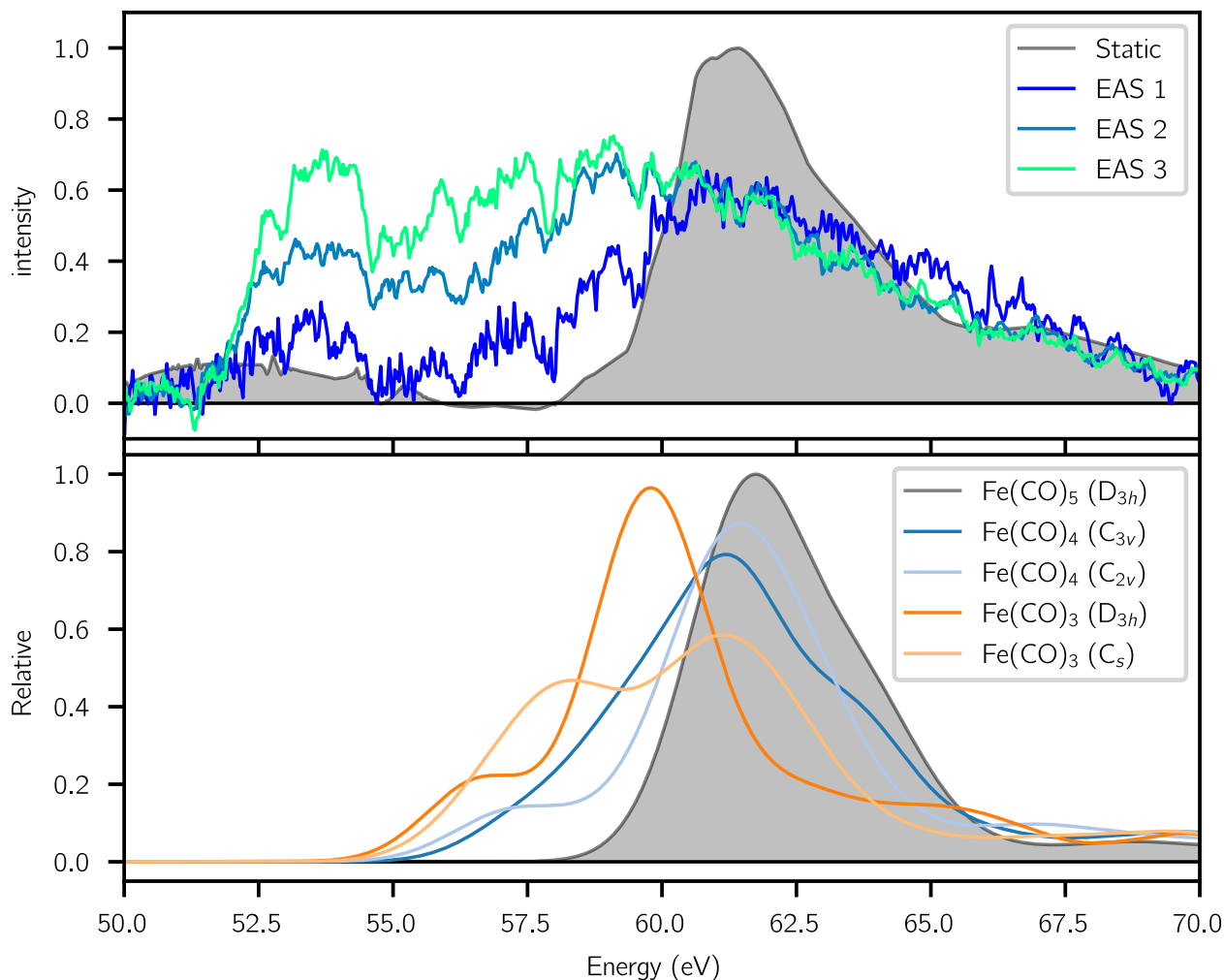
Figure 4.7: Top: evolution-associated spectra (EAS) extracted from the experimental XUV-TRAS. Bottom: EA-TDDFT calculated absorption spectra of unsaturated iron carbonyls.

from different unsaturated iron carbonyls under consideration. For clearer comparison with the EA-TDDFT spectra, they add the static absorption spectrum, at the estimated 9% excitation fraction, to the EADS and generate evolution-associated spectra (EAS) as shown in the left panel of Figure 4.7.

The Figure plots the three EAS that evolve sequentially, which we label from 1 - 3 in increasing order of appearance. EAS-1 appears within the instrument response function and decays with a 120-fs time constant. As discussed in the previous Subsection, we ascribe this to the evolution of the MC excited states during first CO dissociation. As EAS-1 decays, EAS-2 grows in and subsequently decays with a time constant of 3300 fs. Compared to EAS-

1, EAS-2 features a greater absorption between 52 and 60 eV as well as a near absence of intensity at 65 eV. Based on the EA-TDDFT spectra, we propose a significant contribution from $Fe(CO)_4$ to EAS-2. Upon the 3300 fs decay of EAS-2, EAS-3 grows in and does not further evolve within the experimental time window. The dominant characteristic of EAS-3 is a a significant increase in absorption from 52 to 57 eV. However, the strengths of the absorption at 58.5 eV and 61 eV remain constant. In a similar way, the EA-TDDFT spectrum for $Fe(CO)_3$ shows a pronounced absorption at photon energies lower than 56 eV compared to the calculated spectra of $Fe(CO)_4$ and we thus ascribe significant contributions from the final photoproduct to EAS-3. We disregard any meaningful contribution to the spectra from $Fe(CO)_3$ in the D_{3h} geometry as it is roughly 0.4 eV higher in energy than the C_s geometry (Table 4.3).

4.3   Summary

We employed ultrafast XUV-TRAS and electronic structure calculations to characterize excited state evolution and photoproduct formation in the dissociation of $Fe(CO)_5$, a textbook photocatalyst, at 266 nm. Our work is the first to reveal key spectroscopic signatures evolving on a 100-fs timescale due to internal conversion among metal-centered excited-states during first CO loss. Further, this work discovered a minimum energy crossing point for a C_{2v} structure along the first CO dissociation coordinate, which allows for population to funnel into the lowest, closed-shell singlet excited state (S_1). With computations, we found isoenergetic minima of C_{2v} and C_{3v} symmetries for $Fe(CO)_4$ on the S_1 surface, and comparison between the experimental and calculated spectra suggest contributions to the XUV-TRAS from both of them. This is further supported by the small interconversion barrier between the two minima, which leads to rapid structural fluctuations between them. This work discerned the presence of the C_s structure of $Fe(CO)_3$, formed upon thermal dissociation of $Fe(CO)_4$ on the S_1 surface on a 3-ps timescale. This joint experiment-theory work thus represents a significant step forward in our understanding of the ultrafast dynamics of gas-phase $Fe(CO)_5$, afforded by spectroscopic characterization of excited state signatures and lifetimes.

Whereas $Fe(CO)_5$ itself has limited application in the current landscape of organometallic photocatalysis, earth-abundant first-row transition metal-based photocatalysts, including those involving Fe centers, are cost-effective alternatives to rare transition metal counterparts. Our detailed experimental and theoretical characterization of Fe $M_{2,3}$-edge XUV absorption spectroscopy in iron carbonyls serves as a basis for M-edge spectroscopy of other Fe-based photocatalysts in the gas phase which are important because they afford a baseline understanding of excited state dynamics in the absence of environmental influences. Furthermore, these studies also enable application of incisive experimental tools and the use of high-level electronic structure theory that is often difficult to carry out in condensed phase systems. Our work serves as a valuable starting point for future studies of excited-state dynamics in other gas-phase organometallic compounds to aid in efforts towards the ultimate goal of tailoring molecular properties to achieve specific reaction outcomes in photocatalysis.

4.4 Methods

4.4.1 Experimental Methods

4.4.2 Computational Methods

We compute energies at the ωB97M-V/aug-pcseg-3 level of theory, using optimized structures and zero-point energy corrections obtained at the ωB97M-V/aug-pcseg-1 level of theory. The choice of functional was informed by its performance on transition metal thermochemistry in recent benchmarks. [315, 316] As shown in Table 4.3, rCAM-B3LYP and coupled-cluster theory produce similar energies. All TDDFT calculations make use of the Tamm-Dancoff approximation and the LR-TDDFT and EA-TDDFT calculations use the rCAM-B3LYP functional. The S$_2$ / S$_1$ MECP optimizations are carried with spin-flip spin-adapted TDDFT, employing the BHHLYP functional and a collinear approximation to the exchange-correlation kernel. A (99, 590) grid was used for the evaluation of the exchange-correlation potential in DFT calculations. The CASSCF calculations are explained in Section 2.1 of the SI of the paper.

The excited state XUV spectra were calculated with the one-centered non-orthogonal configuration interaction singles theory (1C-NOCIS), described in detail in Chapter 3. [109, 184]

The single point energy calculations using DFT and coupled cluster theory, and the excited state calculations using LR-TDDFT, EA-TDDFT, and 1C-NOCIS account for scalar relativistic effects via the spin-free exact two-component one-electron Hamiltonian (X2C). [147] However, all geometry optimizations, frequency calculations, and CASSCF calculation were carried out without X2C. The amount of spin-orbit-coupling (SOC) present in the M$_{2,3}$-edge is mild, as detailed in the Section 3.4.4 of the SI in the paper, and is disregarded. All presented calculated XUV-TRAS are broadened with a Gaussian of σ =1.0 eV, while EA-TDDFT calculated spectra are shifted by +1.6 eV and 1C-NOCIS calculated spectra are shifted by -1.7 eV to align the calculated spectra with the broad main feature centered at 62 eV with experiment.

Chapter 5

Concluding remarks

In this thesis, we have explored the leveraging of orbital-optimized references for the description of core excited states. In a sense, these references are fine ingredients for that purpose because explicitly relaxing the orbitals in a presence of a core hole confers them the contraction they seek as a result of an electron exciting the core region. On Chapter 2, we focused on using core-hole relaxed references as a starting point for state-specific coupled-cluster calculations of core ionizations and excitations at the K-edge. Therein, we proposed a variety of schemes to address the numerical instabilities and divergences that plague the coupled-cluster equations running on top of non-Aufbau references, as well as the spin-contamination inherent to single-reference theories for low-spin excited states. With a statistical performance on the order of 0.2 RMSE using coupled-cluster singles and doubles for a variety of small main group closed-shell molecules, the accuracy of the method proposed nears experimental uncertainties. On Chapter 3, we moved on to define a zeroth order model for the description of core excited states on top of valence excited states that draws upon the one-centered non-orthogonal configuration interactions singlets theory. Crucially, our 1C-NOCIS model relies on core-hole relaxed orbitals for better accuracy, is spin-pure, and is efficient. This method promises to be of great use for the quickly-developing field of ultra-fast core spectroscopy for the study of valence excited states. In fact, we had the opportunity to test that promise in Chapter 5, where we used 1C-NOCIS to uncover characteristic signatures of the valence excited states of $Fe(CO)_5$ during photodissociation at the iron M-edge. Where to from here?

An interesting finding from Chapter 2 was the relatively poor performance of ΔMP2 for describing core excited states. After sitting on this subject for over a year after publication, I am now of the opinion that the performance of ΔMP2 is truly not bad at all. It still performs much better than traditional response theories! Rather, the performance of ΔHF is impressively good for its sophistication - almost better than it should be - and the fact that ΔMP2 does not deliver the same degree of improvement as MP2 does for HF energies of the ground state, makes it look bad. In fact, a similar relative performance was found by Garner and Neuscamman when comparing ESMF and ESMP2 results for the same purpose. [162] This begs the question as to whether the recent developments in second-order perturbation

theory could offer an improvement. [51–56, 58, 59, 317, 318] Furthermore, the recently proposed excited-state-specific pseudoprojected CC theory (ESPCC) provides a more natural framework for describing the multi-configurational singlet core excited states. [226] Adapting the projector within ESPCC theory to incorporate the schemes presented in Section 2.4 and exploring its performance for core excitations would be exciting.

In regards to 1C-NOCIS 2eOS, the most obvious path forward is to explore a way to include dynamic correlation in the theory and there are two main roads to do so. The first road is to make use of DFT in some manner. Ideally this would be a derivable generalization, such as EA-TDDFT generalized the EA-CIS (namely 1C-NOCIS) theory for closed-shell systems. [114, 267] Alternativley, it could be a theory that simply relies on DFT orbitals within CI formalism such as DFT-MRCI or (s)-TDDFT. [241–243, 250, 251] The second road is to correlate 1C-NOCIS with wave function theory. There is already an abundant number of templates for second-order corrections to CIS excitation energies in the ground state and the first step would be to find a way to transfer these ideas to 1C-NOCIS for closed-shell systems. [319–321]

In general, another exciting avenue for exploration in computational core spectroscopy is going beyond the K-edge. Orbital-specific methods relying on DFT have been tremendously successful for the K-edge, but it may just be that the functional choice is not transferable to the edges beyond. Here, wave function theory could show one of its advantages. While advances in this direction are still not a walk in the park, since this requires inclusion of higher-order relativistic corrections, there is clear progress on the front of spin-orbit coupling for core spectroscopy. [248, 322–326] Furthermore, spin-orbit coupling may be mild for the edges exciting electrons of s character, such as the L_1-, M_1-, ... edges so that a study on this is a ripe fruit ready to be picked.

In regards to the state-specific correlated methods, there is a number of technical developments and benchmarking studies that would be incredibly valuable to enable the new theories presented here to serve their purpose on chemically-relevant systems of ever-increasing size. The convergence of core excitation energies with correlated state-specific methods is slow when using standard correlation-consistent basis designed for ground state chemistry in mind. Thus, designing a correlation-consistent version of the pcX-n series of basis sets - recently designed to serve mean-field orbital-optimized methods for core spectroscopy - would be very useful because the true performance of the correlated methods only shows near the basis set limit. Furthermore, benchmarking the use of locally-dense basis sets as well as custom frozen-core schemes could really bolster the range of applicability of state-specific correlated methods. One last desire of mine here would be the design of specialized windowing algorithms to carry out the appropriate projections required to stabilize the convergence of the coupled-cluster equations on core-excited references.

Finally, we have to use these wonderful tools! While the development of computational core spectroscopy is still on its teenage years, the field is robust enough to use the methods it has designed to explore interesting chemical phenomena. For example, the recent development of efficient methods capable of producing core absorption spectra with sub-eV accuracy for systems on the order of tens of atoms is opening the door to computations to

assist XAS studies on condensed-phase systems. [267, 327] Furthermore, the resolution of experimental ultra-fast core spectroscopy will only continue to improve and just a few dozen of experiments have been carried out in the past decade; the number of interesting systems to study and learn from is vast.

Bibliography

(1) Ezoe, Y.; Ohashi, T.; Mitsuda, K. High-resolution X-ray spectroscopy of astrophysical plasmas with X-ray microcalorimeters. *Reviews of Modern Plasma Physics* **2021**, *5*, 1–43.

(2) Macmillan, S. N.; Lancaster, K. M. X-ray Spectroscopic Interrogation of Transition-Metal-Mediated Homogeneous Catalysis: Primer and Case Studies. *ACS Catalysis* **2017**, *7*, 1776–1791.

(3) Kraus, P. M.; Zürch, M.; Cushing, S. K.; Neumark, D. M.; Leone, S. R. The ultrafast X-ray spectroscopic revolution in chemical dynamics. *Nature Reviews Chemistry* **2018**, *2*, 82–94.

(4) Geneaux, R.; Marroux, H. J.; Guggenmos, A.; Neumark, D. M.; Leone, S. R. Transient absorption spectroscopy using high harmonic generation: a review of ultrafast X-ray dynamics in molecules and solids. *Philosophical Transactions of the Royal Society A* **2019**, *377*, 20170463.

(5) Maiuri, M.; Garavelli, M.; Cerullo, G. Ultrafast Spectroscopy: State of the Art and Open Challenges. *Journal of the American Chemical Society* **2020**, *142*, 3–15.

(6) Summers, A. M. et al. Realizing Attosecond Core-Level X-ray Spectroscopy for the Investigation of Condensed Matter Systems. *Ultrafast Science* **2023**, *3*, 4.

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

(8) Littlejohn, R. G. Quantum Mechanics, Physics 221AB, Lecture Notes, 2021-2022.

(9) Piela, L., *Ideas of Quantum Chemistry*, 2nd ed.; Elsevier: Amsterdam, Netherlands, 2014.

(10) Szabo, A.; Ostlund, N. S., *Modern Quantum Chemistry, Introduction to Advanced Electronic Structure Theory*, 1st ed.; Dover Publications, Inc.: Mineola, New York, 1996.

(11) Schrödinger, E. An Undulatory Theory of the Mechanics of Atoms and Molecules. *Phys. Rev.* **1926**, *28*, 1049–1070.

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

(13)   Pauli, W. Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren. *Zeitschrift für Physik* **1925**, *31*, 765–783.

(14)   Hund, F. Zur Deutung verwickelter Spektren, insbesondere der Elemente Scandium bis Nickel. *Zeitschrift für Physik* **1925**, *33*, 345–371.

(15)   Rubio, M.; Serrano-Andrés, L.; Merchán, M. Excited states of the water molecule: Analysis of the valence and Rydberg character. *Journal of Chemical Physics* **2008**, *128*, 104305.

(16)   Hartree, D. R. The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part I. Theory and Methods. *Mathematical Proceedings of the Cambridge Philosophical Society* **1928**, *24*, 89–110.

(17)   Hartree, D. R. The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part II. Some Results and Discussion. *Mathematical Proceedings of the Cambridge Philosophical Society* **1928**, *24*, 111–132.

(18)   Hartree, D. R. The Wave Mechanics of an Atom with a non-Coulomb Central Field. Part III. Term Values and Intensities in Series in Optical Spectra. *Mathematical Proceedings of the Cambridge Philosophical Society* **1928**, *24*, 426–437.

(19)   Slater, J. C. Note on Hartree's Method. *Physical Review* **1930**, *35*, 210.

(20)   Fock, V. Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems. *Zeitschrift für Physik* **1930**, *61*, 126–148.

(21)   Dirac, P. A. M. On the theory of quantum mechanics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **1926**, *112*, 661–677.

(22)   Heisenberg, W. Mehrkörperproblem und Resonanz in der Quantenmechanik. *Zeitschrift für Physik* **1926**, *38*, 411–426.

(23)   Slater, J. C. The Theory of Complex Spectra. *Physical Review* **1929**, *34*, 1293.

(24)   Fukutome, H. Unrestricted Hartree–Fock theory and its applications to molecules and chemical reactions. *International Journal of Quantum Chemistry* **1981**, *20*, 955–1065.

(25)   Shee, J.; Loipersberger, M.; Hait, D.; Lee, J.; Head-Gordon, M. Revealing the nature of electron correlation in transition metal complexes with symmetry breaking and chemical intuition. *Journal of Chemical Physics* **2021**, *154*, 194109.

(26)   Pople, J. A.; Nesbet, R. K. Self-Consistent Orbitals for Radicals. *The Journal of Chemical Physics* **1954**, *22*, 571.

(27)   Lennard-Jones, J. E. The electronic structure of some diatomic molecules. *Transactions of the Faraday Society* **1929**, *25*, 668–686.

(28)   Hall, G. G. The Lennard-Jones paper of 1929 and the foundations of Molecular Orbital Theory. *Advances in Quantum Chemistry* **1991**, *22*, 1–6.

(29) Boys, S. F. Electronic wave functions - I. A general method of calculation for the stationary states of any molecular system. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **1950**, *200*, 542–554.

(30) Jensen, F. Atomic orbital basis sets. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2013**, *3*, 273–295.

(31) Roothaan, C. C. New Developments in Molecular Orbital Theory. *Reviews of Modern Physics* **1951**, *23*, 69.

(32) Hall, G. G. The molecular orbital theory of chemical valency VIII. A method of calculating ionization potentials. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **1951**, *205*, 541–552.

(33) Lehtola, S. Assessment of Initial Guesses for Self-Consistent Field Calculations. Superposition of Atomic Potentials: Simple yet Efficient. *Journal of Chemical Theory and Computation* **2019**, *15*, 1593–1604.

(34) Pulay, P. Convergence acceleration of iterative sequences. the case of scf iteration. *Chemical Physics Letters* **1980**, *73*, 393–398.

(35) Pulay, P. Improved SCF convergence acceleration. *Journal of Computational Chemistry* **1982**, *3*, 556–560.

(36) Shepard, R.; Minkoff, M. Some comments on the DIIS method. *Molecular Physics* **2007**, *105*, 2839–2848.

(37) Voorhis, T. V.; Head-Gordon, M. A geometric approach to direct minimization. *Molecular Physics* **2002**, *100*, 1713–1721.

(38) Sherrill, C. D.; Schaefer, H. F., *The Configuration Interaction Method: Advances in Highly Correlated Approaches*; Elsevier Masson SAS: 1999; Vol. 34, pp 143–269.

(39) Brillouin, P. L. Les problèmes de perturbations et les champs self-consistents. *Journal de Physique et le Radium* **1932**, *3*, 373–389.

(40) Pauncz, R., *Spin Eigenfunctions*, 1st ed.; Springer: New York, 1979.

(41) Kelly, H. P. Fluorescence yield for the beryllium atom. *Physical Review A* **1974**, *9*, 1582.

(42) Radojević, V.; Johnson, W. R. Multiconfiguration Tamm-Dancoff approximation applied to photoionization of the outer shells of Be and Mg. *Physical Review A* **1985**, *31*, 2991.

(43) Chu, S.; Bovi, D.; Cappelluti, F.; Orellana, A. G.; Martin, H.; Guidoni, L. Effects of Static Correlation between Spin Centers in Multicenter Transition Metal Complexes. *Journal of Chemical Theory and Computation* **2017**, *13*, 4675–4683.

(44) Larsen, H.; Olsen, J.; Jørgensen, P.; Christiansen, O. Full configuration interaction benchmarking of coupled-cluster models for the lowest singlet energy surfaces of N2. *The Journal of Chemical Physics* **2000**, *113*, 6677–6686.

(45) Tew, D. P.; Klopper, W.; Helgaker, T. Electron correlation: The many-body problem at the heart of chemistry. *Journal of Computational Chemistry* **2007**, *28*, 1307–1320.

(46) Coe, J. P.; Paterson, M. J. Investigating Multireference Character and Correlation in Quantum Chemistry. *Journal of Chemical Theory and Computation* **2015**, *11*, 4189–4196.

(47) Stein, C. J.; Reiher, M. Measuring multi-configurational character by orbital entanglement. *Molecular Physics* **2017**, *115*, 2110–2119.

(48) Izsák, R.; Ivanov, A. V.; Blunt, N. S.; Holzmann, N.; Neese, F. Measuring Electron Correlation: The Impact of Symmetry and Orbital Transformations. *Journal of Chemical Theory and Computation* **2023**, *19*, 2703–2720.

(49) Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Physical Review* **1934**, *46*, 618.

(50) Cremer, D.; He, Z. Sixth-order møller-plesset perturbation theory-on the convergence of the mpn series. *Journal of Physical Chemistry* **1996**, *100*, 6173–6188.

(51) Lochan, R. C.; Head-Gordon, M. Orbital-optimized opposite-spin scaled second-order correlation: An economical method to improve the description of open-shell molecules. *Journal of Chemical Physics* **2007**, *126*, 164101.

(52) Neese, F.; Schwabe, T.; Kossmann, S.; Schirmer, B.; Grimme, S. Assessment of orbital-optimized, spin-component scaled second-order many-body perturbation theory for thermochemistry and kinetics. *Journal of Chemical Theory and Computation* **2009**, *5*, 3060–3073.

(53) Rettig, A.; Hait, D.; Bertels, L. W.; Head-Gordon, M. Third-Order Møller-Plesset Theory Made More Useful? The Role of Density Functional Theory Orbitals. *Journal of Chemical Theory and Computation* **2020**, *16*, 7473–7489.

(54) Grimme, S. Improved second-order Møller–Plesset perturbation theory by separate scaling of parallel- and antiparallel-spin pair correlation energies. *The Journal of Chemical Physics* **2003**, *118*, 9095–9102.

(55) Jung, Y.; Lochan, R. C.; Dutoi, A. D.; Head-Gordon, M. Scaled opposite-spin second order Møller–Plesset correlation energy: An economical electronic structure method. *The Journal of Chemical Physics* **2004**, *121*, 9793–9802.

(56) Lochan, R. C.; Jung, Y.; Head-Gordon, M. Scaled opposite spin second order Møller-Plesset theory with improved physical description of long-range dispersion interactions. *Journal of Physical Chemistry A* **2005**, *109*, 7598–7605.

(57) Cremer, D. Møller–Plesset perturbation theory: from small molecule methods to methods for thousands of atoms. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2011**, *1*, 509–530.

(58) Stück, D.; Head-Gordon, M. Regularized orbital-optimized second-order perturbation theory. *Journal of Chemical Physics* **2013**, *139*, 244109.

(59) 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. *Journal of Chemical Theory and Computation* **2018**, *14*, 5203–5219.

(60) Harris, F. E.; Monkhorst, H. J.; Freeman, D. L., *Algebraic and diagrammatic methods in many-fermion theory*, 1st ed.; Oxford Uiversity Press: New York, 1992.

(61) Crawford, T. D.; Scaeffer III, H. F. In Wiley-VCH: Weinheim, Germany, 2009; Chapter An Introduction to Coupled Cluster Theory for Computational Chemists.

(62) Crawford, D. Coupled Cluster Theory, European Summerschoo of Quantum Chemistry, Lecture Notes, 2019.

(63) Scuseria, G. E.; Lee, T. J.; Schaefer, H. F. Accelerating the convergence of the coupled-cluster approach: The use of the DIIS method. *Chemical Physics Letters* **1986**, *130*, 236–239.

(64) Pople, J. A.; Head-Gordon, M.; Raghavachari, K. Quadratic configuration interaction. A general technique for determining electron correlation energies. *The Journal of Chemical Physics* **1987**, *87*, 5968–5975.

(65) Lee, T. J.; Rendell, A. P.; Taylor, P. R. Comparison of the quadratic configuration interaction and coupled-cluster approaches to electron correlation including the effect of triple excitations. *Journal of Physical Chemistry* **1990**, *94*, 5463–5468.

(66) A comparison of the efficiency and accuracy of the quadratic configuration interaction (QCISD), coupled cluster (CCSD), and Brueckner coupled cluster (BCCD) methods. *Chemical Physics Letters* **1992**, *190*, 1–12.

(67) Toulouse, J. Introduction to density-functional theory, European Summerschoo of Quantum Chemistry, Lecture Notes, 2022.

(68) Toulouse, J. Introduction to density-functional theory, Lecture Notes, Lecture Notes, 2022.

(69) Hohenberg, P.; Kohn, W. Inhomogeneous electron gas. *Physical Review* **1964**, *136*, B864.

(70) Kohn, W. v-Representability and Density Functional Theory. *Physical Review Letters* **1983**, *51*, 1596.

(71) Levy, M. Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem. *Proceedings of the National Academy of Sciences* **1979**, *76*, 6062–6065.

(72) Levy, M. Electron densities in search of Hamiltonians. *Physical Review A* **1982**, *26*, 1200.

(73) Kohn, W.; Sham, L. J. Self-consistent equations including exchange and correlation effects. *Physical Review* **1965**, *140*, A1133.

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

(75)   Perdew, J. P.; Ruzsinszky, A.; Tao, J.; Staroverov, V. N.; Scuseria, G. E.; Csonka, G. I. Prescription for the design and selection of density functional approximations: More constraint satisfaction with fewer fits. *Journal of Chemical Physics* **2005**, *123*, 62201.

(76)   Becke, A. D. Density-functional thermochemistry. V. Systematic optimization of exchange-correlation functionals. *The Journal of Chemical Physics* **1997**, *107*, 8554–8560.

(77)   Mardirossian, N.; Head-Gordon, M. How Accurate Are the Minnesota Density Functionals for Noncovalent Interactions, Isomerization Energies, Thermochemistry, and Barrier Heights Involving Molecules Composed of Main-Group Elements? *Journal of Chemical Theory and Computation* **2016**, *12*, 4303–4325.

(78)   Medvedev, M. G.; Bushmarinov, I. S.; Sun, J.; Perdew, J. P.; Lyssenko, K. A. Density functional theory is straying from the path toward the exact functional. *Science* **2017**, *355*, 49–52.

(79)   Cohen, A. J.; Mori-Sánchez, P.; Yang, W. Insights into current limitations of density functional theory. *Science* **2008**, *321*, 792–794.

(80)   Tronc, M.; King, G. C.; Read, F. H. Carbon K-shell excitation in small molecules by high-resolution electron impact. *Journal of Physics B: Atomic and Molecular Physics* **1979**, *12*, 137.

(81)   Hitchcock, A. P.; Brion, C. E. Neon K-shell excitation studied by electron energy-loss spectroscopy. *Journal of Physics B: Atomic and Molecular Physics* **1980**, *13*, 3269–3273.

(82)   Schirmer, J. et al. K-shell excitation of the water, ammonia, and methane molecules using high-resolution photoabsorption spectroscopy. *Physical Review A* **1993**, *47*, 1136–1147.

(83)   Domke, M.; Remmers, G.; Kaindl, G. High-resolution photoionization of core-excited atoms and molecules using synchrotron radiation. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* **1994**, *87*, 173–177.

(84)   Osborne, S. J. et al. The vibrationally resolved participator Auger spectra of selectively excited C 1s(2σ) - $12\pi1$ vibrational states in carbon monoxide. *The Journal of Chemical Physics* **1995**, *102*, 7317–7324.

(85)   Coreno, M. et al. Vibrationally resolved oxygen K$\rightarrow \Pi^*$ spectra of O2 and CO. *Chemical Physics Letters* **1999**, *306*, 269–274.

(86) Prince, K. C.; Avaldi, L.; Coreno, M.; Camilloni, R.; Simone, M. D. Vibrational structure of core to Rydberg state excitations of carbon dioxide and dinitrogen oxide. *Journal of Physics B: Atomic, Molecular and Optical Physics* **1999**, *32*, 2551–2567.

(87) Couto, R. C. et al. The carbon and oxygen K-edge NEXAFS spectra of CO+. *Physical Chemistry Chemical Physics* **2020**, *22*, 16215–16223.

(88) Oversteeg, C. H. V.; Doan, H. Q.; Groot, F. M. D.; Cuk, T. In situ X-ray absorption spectroscopy of transition metal based water oxidation catalysts. *Chemical Society Reviews* **2017**, *46*, 102–125.

(89) Wolf, T. J. et al. Probing ultrafast $\pi\pi^*$ / $n\pi^*$ internal conversion in organic chromophores via K-edge resonant absorption. *Nature Communications* **2017**, *8*, 1–6.

(90) Bhattacherjee, A.; Das Pemmaraju, C.; Schnorr, K.; Attar, A. R.; Leone, S. R. Ultrafast Intersystem Crossing in Acetylacetone via Femtosecond X-ray Transient Absorption at the Carbon K-Edge. *Journal of the American Chemical Society* **2017**, *139*, 16576–16583.

(91) Scutelnic, V. et al. X-ray transient absorption reveals the 1Au ($n\pi^*$) state of pyrazine in electronic relaxation. *Nature Communications* **2021**, *12*, 6–13.

(92) Green, A. E. et al. Spectroscopic Signature of Hydrogen Transfer Dynamics in Acetylacetone. *The International Conference on Ultrafast Phenomena (UP) 2022 (2022), paper Th4A.2* **2022**, Th4A.2.

(93) Ross, A. D. et al. Jahn-Teller distortion and dissociation of CCl4+ by transient X-ray spectroscopy simultaneously at the carbon K- and chlorine L-edge. *Chemical Science* **2022**, *13*, 9310–9320.

(94) Ridente, E. et al. Femtosecond symmetry breaking and coherent relaxation of methane cations via x-ray spectroscopy. *Science* **2023**, *380*, 713–717.

(95) Foresman, J. B.; Head-Gordon, M.; Pople, J. A.; Frisch, M. J. Toward a systematic molecular orbital theory for excited states. *Journal of Physical Chemistry* **1992**, *96*, 135–149.

(96) Davidson, E. R. The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices. *Journal of Computational Physics* **1975**, *17*, 87–94.

(97) Kosugi, N. Modification of the Liu-Davidson method for obtaining one or simultaneously several eigensolutions of a large real-symmetric matrix. *Journal of Computational Physics* **1984**, *55*, 426–436.

(98) Mayer, I. Using singular value decomposition for a compact presentation and improved interpretation of the CIS wave functions. *Chemical Physics Letters* **2007**, *437*, 284–286.

(99)   Vééril, M. et al. QUESTDB: A database of highly accurate excitation energies for the electronic structure community. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2021**, *11*, e1517.

(100) Loos, P. F.; Lipparini, F.; Matthews, D. A.; Blondel, A.; Jacquemin, D. A Mountaineering Strategy to Excited States: Revising Reference Values with EOM-CC4. *Journal of Chemical Theory and Computation* **2022**, *18*, 4418–4427.

(101) Loos, P.-F.; Scemama, A.; Blondel, A.; Garniron, Y.; Caffarel, M.; Jacquemin, D. A Mountaineering Strategy to Excited States: Highly Accurate Reference Energies and Benchmarks. *Journal of Chemical Theory and Computation* **2018**, *14*, 4360–4379.

(102) Loos, P. F.; Lipparini, F.; Boggio-Pasqua, M.; Scemama, A.; Jacquemin, D. A Mountaineering Strategy to Excited States: Highly Accurate Energies and Benchmarks for Medium Sized Molecules. *Journal of Chemical Theory and Computation* **2020**, *16*, 1711–1741.

(103) Loos, P. F.; Scemama, A.; Boggio-Pasqua, M.; Jacquemin, D. Mountaineering Strategy to Excited States: Highly Accurate Energies and Benchmarks for Exotic Molecules and Radicals. *Journal of Chemical Theory and Computation* **2020**, *16*, 3720–3736.

(104) Loos, P. F.; Comin, M.; Blase, X.; Jacquemin, D. Reference Energies for Intramolecular Charge-Transfer Excitations. *Journal of Chemical Theory and Computation* **2021**, *17*, 3666–3686.

(105) Loos, P. F.; Jacquemin, D. A Mountaineering Strategy to Excited States: Highly Accurate Energies and Benchmarks for Bicyclic Systems. *Journal of Physical Chemistry A* **2021**, *125*, 10174–10188.

(106) Liang, J.; Feng, X.; Hait, D.; Head-Gordon, M. Revisiting the Performance of Time-Dependent Density Functional Theory for Electronic Excitations: Assessment of 43 Popular and Recently Developed Functionals from Rungs One to Four. *Journal of Chemical Theory and Computation* **2022**, *18*, 3460–3473.

(107) Cederbaum, L. S.; Domcke, W.; Schirmer, J. Many-body theory of core holes. *Physical Review A* **1980**, *22*, 206.

(108) Stener, M.; Fronzoni, G.; de Simone, M. Time dependent density functional theory of core electrons excitations. *Chemical Physics Letters* **2003**, *373*, 115–123.

(109) Oosterbaan, K. J.; White, A. F.; Head-Gordon, M. Non-orthogonal configuration interaction with single substitutions for the calculation of core-excited states. *Journal of Chemical Physics* **2018**, *149*, DOI: 10.1063/1.5023051.

(110) Runge, E.; Gross, E. K. Density-Functional Theory for Time-Dependent Systems. *Physical Review Letters* **1984**, *52*, 997.

(111) Petersilka, M.; Gossmann, U. J.; Gross, E. K. Excitation Energies from Time-Dependent Density-Functional Theory. *Physical Review Letters* **1996**, *76*, 1212.

(112) Casida, M. E. In *Recent Advances in Density Functional Methods*, Chong, D. P., Ed.; World Scientific: Singapore, 1995, pp 155–192.

(113) Hirata, S.; Head-Gordon, M. Time-dependent density functional theory within the Tamm–Dancoff approximation. *Chemical Physics Letters* **1999**, *314*, 291–299.

(114) 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. *The Journal of Physical Chemistry Letters* **2022**, *13*, 9664–9672.

(115) 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). *Journal of Chemical Physics* **2008**, *129*, 184113.

(116) 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). *Journal of Chemical Physics* **2009**, *131*, 144108.

(117) 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. *Physical Chemistry Chemical Physics* **2009**, *11*, 10350–10358.

(118) 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.

(119) 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+. *Chemical Physics Letters* **2013**, *580*, 179–184.

(120) Besley, N. A. Fast Time-Dependent Density Functional Theory Calculations of the X-ray Absorption Spectroscopy of Large Systems. *Journal of Chemical Theory and Computation* **2016**, *12*, 5018–5025.

(121) Besley, N. A. Density Functional Theory Based Methods for the Calculation of X-ray Spectroscopy. *Accounts of Chemical Research* **2020**, *53*, 1306–1315.

(122) Besley, N. A.; Besley, C. A. N. Modeling of the spectroscopy of core electrons with density functional theory. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2021**, *11*, e1527.

(123) 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. *The Journal of Chemical Physics* **1993**, *98*, 7029–7039.

(124) Bartlett, R. J. Coupled-cluster theory and its equation-of-motion extensions. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2012**, *2*, 126–138.

(125) Nooijen, M.; Bartlett, R. J. A new method for excited states: Similarity transformed equation-of-motion coupled-cluster theory. *The Journal of Chemical Physics* **1997**, *106*, 6441–6448.

(126) Nooijen, M.; Bartlett, R. J. Similarity transformed equation-of-motion coupled-cluster theory: Details, examples, and comparisons. *The Journal of Chemical Physics* **1997**, *107*, 6812–6830.

(127) Epifanovsky, E.; Zuev, D.; Feng, X.; Khistyaev, K.; Shao, Y.; Krylov, A. I. General implementation of the resolution-of-the-identity and Cholesky representations of electron repulsion integrals within coupled-cluster and equation-of-motion methods: Theory and benchmarks. *Journal of Chemical Physics* **2013**, *139*, 134105.

(128) Dutta, A. K.; Neese, F.; Izsák, R. Speeding up equation of motion coupled cluster theory with the chain of spheres approximation. *Journal of Chemical Physics* **2016**, *144*, 34102.

(129) Hohenstein, E. G.; Zhao, Y.; Parrish, R. M.; Martínez, T. J. Rank reduced coupled cluster theory. II. Equation-of-motion coupled-cluster singles and doubles. *Journal of Chemical Physics* **2019**, *151*, 164121.

(130) Kucharski, S. A.; Włoch, M.; Musiał, M.; Bartlett, R. J. Coupled-cluster theory for excited electronic states: The full equation-of-motion coupled-cluster single, double, and triple excitation method. *The Journal of Chemical Physics* **2001**, *115*, 8263–8266.

(131) Hirata, S. Higher-order equation-of-motion coupled-cluster methods. *The Journal of Chemical Physics* **2004**, *121*, 51–59.

(132) Coriani, S.; Christiansen, O.; Fransson, T.; Norman, P. Coupled-cluster response theory for near-edge x-ray-absorption fine structure of atoms and molecules. *Physical Review A* **2012**, *85*, 22507.

(133) Kauczor, J.; Norman, P.; Christiansen, O.; Coriani, S. Communication: A reduced-space algorithm for the solution of the complex linear response equations used in coupled cluster damped response theory. *The Journal of Chemical Physics* **2013**, *139*, 211102.

(134) Zuev, D.; Vecharynski, E.; Yang, C.; Orms, N.; Krylov, A. I. New algorithms for iterative matrix-free eigensolvers in quantum chemistry. *Journal of Computational Chemistry* **2015**, *36*, 273–284.

(135) Coriani, S.; Koch, H. Communication: X-ray absorption spectra and core-ionization potentials within a core-valence separated coupled cluster framework. *The Journal of Chemical Physics* **2015**, *143*, 181103.

(136) 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. *Journal of Chemical Theory and Computation* **2019**, *15*, 3117–3133.

(137) Barth, A.; Schirmer, J. Theoretical core-level excitation spectra of N2 and CO by a new polarisation propagator method. *Journal of Physics B: Atomic and Molecular Physics* **1985**, *18*, 867–885.

(138) Liu, J.; Matthews, D.; Coriani, S.; Cheng, L. Benchmark Calculations of K-Edge Ionization Energies for First-Row Elements Using Scalar-Relativistic Core-Valence-Separated Equation-of-Motion Coupled-Cluster Methods. *Journal of Chemical Theory and Computation* **2019**, *15*, 1642–1651.

(139) Faber, R.; Kjønstad, E. F.; Koch, H.; Coriani, S. Spin adapted implementation of EOM-CCSD for triplet excited states: Probing intersystem crossings of acetylacetone at the carbon and oxygen K-edges. *Journal of Chemical Physics* **2019**, *151*, 144107.

(140) Gilbert, A. T.; Besley, N. A.; Gill, P. M. Self-consistent field calculations of excited states using the maximum overlap method (MOM). *Journal of Physical Chemistry A* **2008**, *112*, 13164–13171.

(141) Barca, G. M.; Gilbert, A. T.; Gill, P. M. Simple Models for Difficult Electronic Excitations. *Journal of Chemical Theory and Computation* **2018**, *14*, 1501–1509.

(142) 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. *Journal of Chemical Theory and Computation* **2020**, *16*, 1699–1710.

(143) 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. *Journal of Chemical Theory and Computation* **2020**, *16*, 5067–5082.

(144) Yamaguchi, K.; Jensen, F.; Dorigo, A.; Houk, K. A spin correction procedure for unrestricted Hartree-Fock and Möller-Plesset wavefunctions for singlet diradicals and polyradicals. *Chem. Phys. Lett.* **1988**, *149*, 537–542.

(145) Molecular dynamics in low-spin excited states. *The Journal of Chemical Physics* **1998**, *108*, 4060–4069.

(146) 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. *Journal of Physical Chemistry Letters* **2020**, *11*, 775–786.

(147) Saue, T. Relativistic Hamiltonians for Chemistry: A Primer. *ChemPhysChem* **2011**, *12*, 3077–3094.

(148) Cunha, L. A.; Hait, D.; Kang, R.; Mao, Y.; Head-Gordon, M. Relativistic Orbital-Optimized Density Functional Theory for Accurate Core-Level Spectroscopy. *Journal of Physical Chemistry Letters* **2022**, *13*, 3438–3449.

(149) 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. *Journal of Chemical Physics* **2020**, *153*, 134108.

(150) Hait, D.; Head-Gordon, M. Orbital Optimized Density Functional Theory for Electronic Excited States. *Journal of Physical Chemistry Letters* **2021**, *12*, 4517–4529.

(151) Cembran, A.; Song, L.; Mo, Y.; Gao, J. Block-localized density functional theory (BLDFT), diabatic coupling, and their use in valence Bond theory for representing reactive potential energy surfaces. *Journal of Chemical Theory and Computation* **2009**, *5*, 2702–2716.

(152) Ren, H.; Provorse, M. R.; Bao, P.; Qu, Z.; Gao, J. Multistate Density Functional Theory for Effective Diabatic Electronic Coupling. *Journal of Physical Chemistry Letters* **2016**, *7*, 2286–2293.

(153) Grofe, A.; Chen, X.; Liu, W.; Gao, J. Spin-multiplet components and energy splittings by multistate density functional theory. *Journal of Physical Chemistry Letters* **2017**, *8*, 4838–4845.

(154) Zhao, R. et al. Dynamic-then-Static Approach for Core Excitations of Open-Shell Molecules. *Journal of Physical Chemistry Letters* **2021**, *12*, 7409–7417.

(155) Lu, Y.; Gao, J. Multistate Density Functional Theory of Excited States. *Journal of Physical Chemistry Letters* **2022**, *13*, 7762–7769.

(156) Choi, J. H.; Lebeda, C. F.; Messmer, R. P. Variational principle for excited states: Exact formulation and other extensions. *Chemical Physics Letters* **1970**, *5*, 503–506.

(157) Zhao, L.; Neuscamman, E. An Efficient Variational Principle for the Direct Optimization of Excited States. *Journal of Chemical Theory and Computation* **2016**, *12*, 3436–3440.

(158) Shea, J. A.; Neuscamman, E. Size Consistent Excited States via Algorithmic Transformations between Variational Principles. *Journal of Chemical Theory and Computation* **2017**, *13*, 6078–6088.

(159) Robinson, P. J.; Flores, S. D. P.; Neuscamman, E. Excitation variance matching with limited configuration interaction expansions in variational Monte Carlo. *Journal of Chemical Physics* **2017**, *147*, 164114.

(160) Shea, J. A.; Neuscamman, E. Communication: A mean field platform for excited state quantum chemistry. *Journal of Chemical Physics* **2018**, *149*, 081101.

(161) Kowalczyk, T.; Tsuchimochi, T.; Chen, P. T.; Top, L.; Voorhis, T. V. Excitation energies and Stokes shifts from a restricted open-shell Kohn-Sham approach. *Journal of Chemical Physics* **2013**, *138*, 164101.

(162) Garner, S. M.; Neuscamman, E. Core excitations with excited state mean field and perturbation theory. *Journal of Chemical Physics* **2020**, *153*, 154102.

(163) Ehlert, C.; Klamroth, T. PSIXAS: A Psi4 plugin for efficient simulations of X-ray absorption spectra based on the transition-potential and Δ-Kohn–Sham method. *Journal of Computational Chemistry* **2020**, *41*, 1781–1789.

(164) Meissner, L.; Balkova, A.; Bartlett, R. J. Multiple solutions of the single-reference coupled-cluster method. *Chemical Physics Letters* **1993**, *2*, 12.

(165) S. William J. Hunt; Goddard, W. A. Excited States of H2O using improved virtual orbitals. *Chemical Physics Letters* **1969**, *3*, 414–418.

(166) Ågren, H.; Carravetta, V.; Vahtras, O.; Pettersson, L. G. Direct, atomic orbital, static exchange calculations of photoabsorption spectra of large molecules and clusters. *Chemical Physics Letters* **1994**, *222*, 75–81.

(167) Ågren, H.; Carravetta, V.; Vahtras, O.; Pettersson, L. G. Direct SCF direct static-exchange calculations of electronic spectra. *Theoretical Chemistry Accounts* **1997**, *97*, 14–40.

(168) Carravetta, V.; Plashkevych, O.; Ågren, H. A screened static-exchange potential for core eclectron excitations. *Chemical Physics* **2001**, *263*, 231–242.

(169) Brumboiu, I. E.; Fransson, T. Core–hole delocalization for modeling x-ray spectroscopies: A cautionary tale. *The Journal of Chemical Physics* **2022**, *156*, 214109.

(170) 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. *Journal of Chemical Theory and Computation* **2019**, *15*, 2966–2973.

(171) Carravetta, V.; Couto, R. C.; Ågren, H. X-ray absorption of molecular cations—a new challenge for electronic structure theory. *Journal of Physics: Condensed Matter* **2022**, *34*, 363002.

(172) Oosterbaan, K. J.; White, A. F.; Hait, D.; Head-Gordon, M. Generalized single excitation configuration interaction: An investigation into the impact of the inclusion of non-orthogonality on the calculation of core-excited states. *Physical Chemistry Chemical Physics* **2020**, *22*, 8182–8192.

(173) Hait, D.; Oosterbaan, K. J.; Carter-Fenk, K.; Head-Gordon, M. Computing x-ray absorption spectra from linear-response particles atop optimized holes. *Journal of Chemical Physics* **2022**, *156*, 201104.

(174) Krylov, A. I. Equation-of-motion coupled-cluster methods for open-shell and electronically excited species: the Hitchhiker's guide to Fock space. *Annual review of physical chemistry* **2008**, *59*, 433–462.

(175) Casanova, D.; Krylov, A. I. Spin-flip methods in quantum chemistry. *Physical Chemistry Chemical Physics* **2020**, *22*, 4326–4342.

(176) Park, W.; Alías-Rodríguez, M.; Cho, D.; Lee, S.; Huix-Rotllant, M.; Choi, C. H. Mixed-Reference Spin-Flip Time-Dependent Density Functional Theory for Accurate X-ray Absorption Spectroscopy. *Journal of Chemical Theory and Computation* **2022**, *18*, 6240–6250.

(177) Triguero, L.; Pettersson, L.; 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. *Physical Review B - Condensed Matter and Materials Physics* **1998**, *58*, 8097–8110.

(178) Triguero, L.; Plashkevych, O.; Pettersson, L. G.; Ågren, H. Separate state vs. transition state Kohn-Sham calculations of X-ray photoelectron binding energies and chemical shifts. *Journal of Electron Spectroscopy and Related Phenomena* **1999**, *104*, 195–207.

(179) Jana, S.; Herbert, J. M. Slater transition methods for core-level electron binding energies. *Journal of Chemical Physics* **2023**, *158*, 94111.

(180) Jana, S.; Herbert, J. M. Fractional-Electron and Transition-Potential Methods for Core-to-Valence Excitation Energies Using Density Functional Theory. *Journal of Chemical Theory and Computation* **2023**, *19*, 4100–4113.

(181) Simons, M.; Matthews, D. A. Transition-potential coupled cluster. *Journal of Chemical Physics* **2021**, *154*, 14106.

(182) Simons, M.; Matthews, D. A. Transition-potential coupled cluster II: optimisation of the core orbital occupation number. *Molecular Physics* **2023**, *121*, 11–12.

(183) 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. *Physical Chemistry Chemical Physics* **2022**, *24*, 20728–20741.

(184) Arias-Martinez, J. E.; Wu, H.; Head-Gordon, M. Generalization of one-center non orthogonal configuration interaction singles to open shell singlet reference states: Theory and application to valence-core pump-probe states in acetylacetone. *arXiv e-prints* **2023**, arXiv:2310.09684.

(185) Troß, J. et al. Femtosecond core-level spectroscopy reveals signatures of transient excited states in the photodissociation of $Fe(CO)_5$. *ChemRxiv e-prints* **2023**, 10.26434/chemrxiv-2023–m1ggn.

(186) Talbot, J. J.; Arias-Martinez, J. E.; Cotton, S. J.; Head-Gordon, M. Fantastical excited state optimized structures and where to find them. *The Journal of Chemical Physics* **2023**, *159*, 171102.

(187) Besley, N. A.; Gilbert, A. T.; Gill, P. M. Self-consistent-field calculations of core excited states. *The Journal of Chemical Physics* **2009**, *130*, 124308.

(188) Duflot, D.; Flament, J. P.; Heinesch, J.; Hubin-Franskin, M. J. The K-shell spectra of tetrahydrofuran studied by electron energy loss spectroscopy and ab initio calculations. *Chemical Physics Letters* **2010**, *495*, 27–32.

(189) Shim, J.; Klobukowski, M.; Barysz, M.; Leszczynski, J. Calibration and applications of the ΔmP2 method for calculating core electron binding energies. *Physical Chemistry Chemical Physics* **2011**, *13*, 5703–5711.

(190) Ljubić, I. Reliability of density functional and perturbation theories for calculating core-ionization spectra of free radicals. *Journal of Chemical Theory and Computation* **2014**, *10*, 2333–2343.

(191) Zheng, X.; Cheng, L. Performance of Delta-Coupled-Cluster Methods for Calculations of Core-Ionization Energies of First-Row Elements. *Journal of Chemical Theory and Computation* **2019**, *15*, 4945–4955.

(192) 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. *The Journal of Chemical Physics* **2019**, *151*, 214103.

(193) 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. *Journal of Physical Chemistry A* **2020**, *124*, 4413–4426.

(194) Zheng, X.; Zhang, C.; Jin, Z.; Southworth, S. H.; Cheng, L. Benchmark relativistic delta-coupled-cluster calculations of K-edge core-ionization energies of third-row elements. *Physical Chemistry Chemical Physics* **2022**, *24*, 13587–13596.

(195) Matthews, D. A. EOM-CC methods with approximate triple excitations applied to core excitation and ionisation energies. *Molecular Physics* **2020**, *118*, e1771448.

(196) Stanton, J. F.; Gauss, J.; Watts, J. D.; Bartlett, R. J. A direct product decomposition approach for symmetry exploitation in many-body methods. I. Energy calculations. *The Journal of Chemical Physics* **1991**, *94*, 4334–4345.

(197) Kucharski, S. A.; Bartlett, R. J. Hilbert space multireference coupled-cluster methods. I. the single and double excitation model. *Journal of Chemical Physics* **1991**, *95*, 8227–8238.

(198) Balková, A.; Bartlett, R. J. Coupled-cluster method for open-shell singlet states. *Chemical Physics Letters* **1992**, *193*, 364–372.

(199) Nooijen, M.; Bartlett, R. J. Description of core-excitation spectra by the open-shell electron-attachment equation-of-motion coupled cluster method. *The Journal of Chemical Physics* **1995**, *102*, 6735–6756.

(200) *NIST Computational Chemistry Comparison and Benchmark Database, NIST Standard Reference Database Number 101*; tech. rep. Federal Information Processing Standards Publications (FIPS PUBS) 140-2, Change Notice 2 December 03, 2002; Washington, D.C.: U.S. Department of Commerce, 2022.

(201) Takahashi, O. Relativistic corrections for single- and double-core excitation at the K- and L-edges from Li to Kr. *Computational and Theoretical Chemistry* **2017**, *1102*, 80–86.

(202) Ambroise, M. A.; Jensen, F. Probing Basis Set Requirements for Calculating Core Ionization and Core Excitation Spectroscopy by the *delta* Self-Consistent-Field Approach. *Journal of Chemical Theory and Computation* **2019**, *15*, 325–337.

(203) Ambroise, M. A.; Dreuw, A.; Jensen, F. Probing Basis Set Requirements for Calculating Core Ionization and Core Excitation Spectra Using Correlated Wave Function Methods. *Journal of Chemical Theory and Computation* **2021**, *17*, 2832–2842.

(204) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. Basis-set convergence of correlated calculations on water. *Journal of Chemical Physics* **1997**, *106*, 9639–9646.

(205) Halkier, A. et al. Basis-set convergence in correlated calculations on Ne, N2, and H2O. *Chemical Physics Letters* **1998**, *286*, 243–252.

(206) 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*, 1st ed.; Elsevier Inc.: 2019; Vol. 79, pp 241–261.

(207) 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.

(208) Peterson, K. A.; Dunning Jr, T. H. Accurate correlation consistent basis sets for molecular core–valence correlation effects: The second row atoms Al–Ar, and the first row atoms B–Ne revisited. *J. Chem. Phys.* **2002**, *117*, 10548–10560.

(209) Pireaux, J. J. et al. Core-electron relaxation energies and valence-band formation of linear alkanes studied in the gas phase by means of electron spectroscopy. *Physical Review A* **1976**, *14*, 2133.

(210) Jolly, W. L.; Bomben, K. D.; Eyermann, C. J. Core-electron binding energies for gaseous atoms and molecules. *Atomic Data and Nuclear Data Tables* **1984**, *31*, 433–493.

(211) Oliphant, N.; Adamowicz, L. Multireference coupled-cluster method using a single-reference formalism. *The Journal of Chemical Physics* **1991**, *94*, 1229.

(212) Sun, J.; Ruzsinszky, A.; Perdew, J. Strongly Constrained and Appropriately Normed Semilocal Density Functional. *Physical Review Letters* **2015**, *115*, 036402.

(213) Kramida, A.; Martin, W. C. A Compilation of Energy Levels and Wavelengths for the Spectrum of Neutral Beryllium (Be I). *Journal of Physical and Chemical Reference Data* **2009**, *26*, 1185.

(214) Shaw, D. A.; King, G. C.; Read, F. H.; Cvejanovic, D. The observation of electric-dipole-forbidden inner-shell transitions in N2 and Ar by the electron energy-loss technique. *Journal of Physics B: Atomic and Molecular Physics* **1982**, *15*, 1785–1793.

(215) Hitchcock, A. P.; Brion, C. E. Carbon K-shell excitation of C2H2, C2H4, C2H6 and C6H6 by 2.5 keV electron impact. *Journal of Electron Spectroscopy and Related Phenomena* **1977**, *10*, 317–330.

(216) Remmers, G. et al. High-resolution K-shell photoabsorption in formaldehyde. *Physical Review A* **1992**, *46*, 3935–3944.

(217) Hitchcock, A. P.; Brion, C. E. Inner shell electron energy loss studies of HCN and C2N2. *Chemical Physics* **1979**, *37*, 319–331.

(218) Sodhi, R. N.; Brion, C. E. Reference energies for inner shell electron energy-loss spectroscopy. *Journal of Electron Spectroscopy and Related Phenomena* **1984**, *34*, 363–372.

(219) Prince, K. C.; Richter, R.; Simone, M. D.; Alagia, M.; Coreno, M. Near edge x-ray absorption spectra of some small polyatomic molecules. *Journal of Physical Chemistry A* **2003**, *107*, 1955–1963.

(220) Hitchcock, A. P.; Brion, C. E. K-shell excitation of HF and F2 studied by electron energy-loss spectroscopy. *Journal of Physics B: Atomic and Molecular Physics* **1981**, *14*, 4399–4413.

(221) Müller, A. et al. Photoionization of Ne Atoms and Ne + Ions Near the K Edge: PrecisionSpectroscopy and Absolute Cross-sections. *The Astrophysical Journal* **2017**, *836*, 166.

(222) Hempelmann, A.; Piancastelli, M. N.; Heiser, F.; Gessner, O.; Rüdel, A.; Becker, U. Resonant photofragmentation of methanol at the carbon and oxygen K-edge by high-resolution ion-yield spectroscopy. *Journal of Physics B: Atomic, Molecular and Optical Physics* **1999**, *32*, 2677–2689.

(223) Carravetta, V.; Ågren, H.; Cesar, A. Is the core photoelectron band of neon asymmetric or not? *Chemical Physics Letters* **1991**, *180*, 358–364.

(224) Simone, M. D.; Coreno, M.; Alagia, M.; Richter, R.; Prince, K. C. Inner shell excitation spectroscopy of the tetrahedral molecules CX4 (X = H, F, Cl). *Journal of Physics B: Atomic, Molecular and Optical Physics* **2002**, *35*, 61–75.

(225) Duflot, D.; Flament, J. P.; Giuliani, A.; Heinesch, J.; Hubin-Franskin, M. J. Core shell excitation of furan at the O1s and C1s edges: An experimental and ab initio study. *The Journal of Chemical Physics* **2003**, *119*, 8946.

(226) Tuckman, H.; Neuscamman, E. Excited-State-Specific Pseudoprojected Coupled-Cluster Theory. *Journal of Chemical Theory and Computation* **2023**, *19*, 6171.

(227) Chergui, M.; Beye, M.; Mukamel, S.; Svetina, C.; Masciovecchio, C. Progress and prospects in nonlinear extreme-ultraviolet and X-ray optics and spectroscopy. *Nature Reviews Physics 2023 5:10* **2023**, *5*, 578–596.

(228) Bengtsson, S. et al. Space-time control of free induction decay in the extreme ultraviolet. *Nature Photonics* **2017**, *11*, 252–258.

(229) Barreau, L.; Ross, A. D.; Garg, S.; Kraus, P. M.; Neumark, D. M.; Leone, S. R. Efficient table-top dual-wavelength beamline for ultrafast transient absorption spectroscopy in the soft X-ray region. *Scientific Reports* **2020**, *10*, 1–9.

(230) 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. *Journal of Chemical Physics* **2013**, *138*, 204101.

(231) Tsuru, S.; Vidal, M. L.; Pápai, M.; Krylov, A. I.; Møller, K. B.; Coriani, S. An assessment of different electronic structure approaches for modeling time-resolved X-ray absorption spectroscopy. *Structural Dynamics* **2021**, *8*, 24101.

(232) Loos, P. F.; Boggio-Pasqua, M.; Scemama, A.; Caffarel, M.; Jacquemin, D. Reference Energies for Double Excitations. *Journal of Chemical Theory and Computation* **2019**, *15*, 1939–1956.

(233) Ruberti, M. et al. Total photoionization cross-sections of excited electronic states by the algebraic diagrammatic construction-Stieltjes-Lanczos method. *Journal of Chemical Physics* **2014**, *140*, 184107.

(234) Leitner, J.; Dempwolff, A. L.; Dreuw, A. The fourth-order algebraic diagrammatic construction scheme for the polarization propagator. *Journal of Chemical Physics* **2022**, *157*, 184101.

(235) Neville, S. P. et al. Excited state X-ray absorption spectroscopy: Probing both electronic and structural dynamics. *The Journal of Chemical Physics* **2016**, *145*, 144307.

(236) Neville, S. P. et al. Beyond structure: ultrafast X-ray absorption spectroscopy as a probe of non-adiabatic wavepacket dynamics. *Faraday Discussions* **2016**, *194*, 117–145.

(237) Neville, S. P.; Chergui, M.; Stolow, A.; Schuurman, M. S. Ultrafast X-Ray Spectroscopy of Conical Intersections. *Physical Review Letters* **2018**, *120*, 243001.

(238) Kaczun, T.; Dempwolff, A. L.; Huang, X.; Gelin, M. F.; Domcke, W.; Dreuw, A. Tuning UV Pump X-ray Probe Spectroscopy on the Nitrogen K Edge Reveals the Radiationless Relaxation of Pyrazine: Ab Initio Simulations Using the Quasiclassical Doorway-Window Approximation. *Journal of Physical Chemistry Letters* **2023**, *14*, 5648–5656.

(239) Hohenstein, E. G. et al. Predictions of Pre-edge Features in Time-Resolved Near-Edge X-ray Absorption Fine Structure Spectroscopy from Hole-Hole Tamm-Dancoff-Approximated Density Functional Theory. *Journal of Chemical Theory and Computation* **2021**, *17*, 7120–7133.

(240) Attar, A. R. et al. Femtosecond x-ray spectroscopy of an electrocyclic ring-opening reaction. *Science* **2017**, *356*, 54–59.

(241) Li, Z.; Liu, W. Spin-adapted open-shell random phase approximation and time-dependent density functional theory. I. Theory. *Journal of Chemical Physics* **2010**, *133*, 64106.

(242) Li, Z.; Liu, W.; Zhang, Y.; Suo, B. Spin-adapted open-shell time-dependent density functional theory. II. Theory and pilot application. *Journal of Chemical Physics* **2011**, *134*, 134101.

(243) Li, Z.; Liu, W. Spin-adapted open-shell time-dependent density functional theory. III. An even better and simpler formulation. *Journal of Chemical Physics* **2011**, *135*, 194106.

(244) Haugen, E. A.; Hait, D.; Scutelnic, V.; Xue, T.; Head-Gordon, M.; Leone, S. R. Ultrafast X-ray Spectroscopy of Intersystem Crossing in Hexafluoroacetylacetone: Chromophore Photophysics and Spectral Changes in the Face of Electron-Withdrawing Groups. *Journal of Physical Chemistry A* **2023**, *127*, 634–644.

(245) Garner, S. M.; Haugen, E. A.; Leone, S. R.; Neuscamman, E. Spin Coupling Effect on Geometry-Dependent X-ray Absorption of Diradicals, 2023.

(246) Li, X.; Govind, N.; Isborn, C.; Deprince, A. E.; Lopata, K. Real-Time Time-Dependent Electronic Structure Theory. *Chemical Reviews* **2020**, *120*, 9951–9993.

(247) Chen, M.; Lopata, K. First-Principles Simulations of X-ray Transient Absorption for Probing Attosecond Electron Dynamics. *Journal of Chemical Theory and Computation* **2020**, *16*, 4470–4478.

(248) Moitra, T.; Konecny, L.; Kadek, M.; Rubio, A.; Repisky, M. Accurate Relativistic Real-Time Time-Dependent Density Functional Theory for Valence and Core Attosecond Transient Absorption Spectroscopy. *The Journal of Physical Chemistry Letters* **2023**, *14*, 1714–1724.

(249) Herbert, J. M.; Zhu, Y.; Alam, B.; Ojha, A. K. Time-Dependent Density Functional Theory for X-ray Absorption Spectra: Comparing the Real-Time Approach to Linear Response. *Journal of Chemical Theory and Computation* **2023**, DOI: `10.1021/ACS.JCTC.3C00673`.

(250) Grimme, S.; Waletzke, M. A combination of Kohn-Sham density functional theory and multi-reference configuration interaction methods. *Journal of Chemical Physics* **1999**, *111*, 5645–5655.

(251) Seidu, I.; Neville, S. P.; Macdonell, R. J.; Schuurman, M. S. Resolving competing conical intersection pathways: time-resolved X-ray absorption spectroscopy of trans-1,3-butadiene. *Physical Chemistry Chemical Physics* **2022**, *24*, 1345–1354.

(252) List, N. H.; Dempwolff, A. L.; Dreuw, A.; Norman, P.; Martínez, T. J. Probing competing relaxation pathways in malonaldehyde with transient X-ray absorption spectroscopy. *Chemical Science* **2020**, *11*, 4180–4193.

(253) Carlini, L. et al. Electron and ion spectroscopy of azobenzene in the valence and core shells. *Journal of Chemical Physics* **2023**, *158*, 054201.

(254) Huang, M.; Evangelista, F. A. A study of core-excited states of organic molecules computed with the generalized active space driven similarity renormalization group. *Journal of Chemical Physics* **2023**, *158*, 124112.

(255) Marin, K.; Huang, M.; Evangelista, F. A. Signatures of diradicals in x-ray absorption spectroscopy. *Journal of Chemical Physics* **2023**, *158*, 151101.

(256) Mazin, I. M.; Sokolov, A. Y. Core-Excited States and X-ray Absorption Spectra from Multireference Algebraic Diagrammatic Construction Theory. *Journal of Chemical Theory and Computation* **2023**, *19*, 4991–5006.

(257) Kossoski, F.; Loos, P. F. State-Specific Configuration Interaction for Excited States. *Journal of Chemical Theory and Computation* **2023**, *19*, 2258–2269.

(258) Kossoski, F.; Loos, P. F. Seniority and Hierarchy Configuration Interaction for Radicals and Excited States. *Journal of Chemical Theory and Computation* **2023**, *19*, 8670.

(259) Schmerwitz, Y. L.; Levi, G.; J'onsson, H. Calculations of Excited Electronic States by Converging on Saddle Points Using Generalized Mode Following. *Journal of Chemical Theory and Computation* **2023**, *19*, 3634–3651.

(260) Epifanovsky, E. et al. Software for the frontiers of quantum chemistry: An overview of developments in the Q-Chem 5 package. *The Journal of Chemical Physics* **2021**, *155*, 084801:1–59.

(261) Jensen, F. Unifying general and segmented contracted basis sets. segmented polarization consistent basis sets. *Journal of Chemical Theory and Computation* **2014**, *10*, 1074–1085.

(262) Saue, T. Relativistic Hamiltonians for Chemistry: A Primer. *ChemPhysChem* **2011**, *12*, 3077–3094.

(263) Subotnik, J. E.; Shenvi, N. A new approach to decoherence and momentum rescaling in the surface hopping algorithm. *Journal of Chemical Physics* **2011**, *134*, 024105.

(264) Moitra, T.; Coriani, S.; Tenorio, B. N. C. Inner-shell photoabsorption and photoionisation cross-sections of valence excited states from asymmetric-Lanczos equation-of-motion coupled cluster singles and doubles theory. *Molecular Physics* **2021**, *119*, 184106.

(265) Lindblad, R. et al. X-Ray Absorption Spectrum of the N2+ Molecular Ion. *Physical Review Letters* **2020**, *124*, 203001.

(266) Squibb, R. J. et al. Acetylacetone photodynamics at a seeded free-electron laser. *Nature Communications* **2018**, *9*, 1–7.

(267) 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. *Physical Chemistry Chemical Physics* **2022**, *24*, 26170–26179.

(268) Chang, K. F. et al. Conical intersection and coherent vibrational dynamics in alkyl iodides captured by attosecond transient absorption spectroscopy. *Journal of Chemical Physics* **2022**, *156*, 114304.

(269) Carter-Fenk, J. T. K. et al. Excited-State Dynamics during Primary C–I Homolysis in Acetyl Iodide Revealed by Ultrafast Core-Level Spectroscopy. *The Journal of Physical Chemistry A* **2023**, *127*, 4103–4114.

(270) Toulson, B. W. et al. Probing C-I bond fission in the UV photochemistry of 2-iodothiophene with core-to-valence transient absorption spectroscopy. *Journal of Chemical Physics* **2023**, *159*, 34304.

(271) Hennig, H. Homogeneous photo catalysis by transition metal complexes. *Coordination Chemistry Reviews* **1999**, *182*, 101–123.

(272) Malik, J. A.; Madani, A.; Pieber, B.; Seeberger, P. H. Evidence for Photocatalyst Involvement in Oxidative Additions of Nickel-Catalyzed Carboxylate *O*-Arylations. *Journal of the American Chemical Society* **2020**, *142*, 11042–11049.

(273) Dumur, F. Recent Advances on Visible Light Metal-Based Photocatalysts for Polymerization under Low Light Intensity. *Catalysts 2019, Vol. 9, Page 736* **2019**, *9*, 736.

(274) *Visible light photocatalysis in organic chemistry*; Stephenson, C. R., Yoon, T. P., MacMillan, D. W., Eds.; Wiley-VCH: Weinheim, Germany, 2018.

(275) Zhang, W. et al. Tracking excited-state charge and spin dynamics in iron coordination complexes. *Nature* **2014**, *509*, 345–348.

(276) Shimbayashi, T.; Fujita, K.-i. Recent advances in homogeneous catalysis via metal–ligand cooperation involving aromatization and dearomatization. *Catalysts* **2020**, *10*, 635.

(277) Terrett, J. A.; Cuthbertson, J. D.; Shurtleff, V. W.; MacMillan, D. W. Switching on elusive organometallic mechanisms with photoredox catalysis. *Nature* **2015**, *524*, 330–334.

(278) Prier, C. K.; Rankic, D. A.; MacMillan, D. W. C. Visible Light Photoredox Catalysis with Transition Metal Complexes: Applications in Organic Synthesis. *Chemical Reviews* **2013**, *113*, 5322–5363.

(279) Williams, G.; Seger, B.; Kamat, P. V. TiO2-Graphene Nanocomposites. UV-Assisted Photocatalytic Reduction of Graphene Oxide. *ACS Nano* **2008**, *2*, 1487–1491.

(280) Arias-Rotondo, D. M.; McCusker, J. K. The photophysics of photoredox catalysis: a roadmap for catalyst design. *Chem. Soc. Rev.* **2016**, *45*, 5803–5820.

(281) Na, H.; Mirica, L. M. Deciphering the mechanism of the Ni-photocatalyzed C–O cross-coupling reaction using a tridentate pyridinophane ligand. *Nature communications* **2022**, *13*, 1–11.

(282) Wrighton, M. Photochemistry of metal carbonyls. *Chemical Reviews* **1974**, *74*, 401–430.

(283) Frenking, G.; Fernández, I.; Holzmann, N.; Pan, S.; Krossing, I.; Zhou, M. Metal–CO Bonding in Mononuclear Transition Metal Carbonyl Complexes. *JACS Au* **2021**, *1*, 623–645.

(284) Cole-Filipiak, N. C.; Troß, J.; Schrader, P.; McCaslin, L. M.; Ramasesha, K. Ultrafast infrared transient absorption spectroscopy of gas-phase $Ni(CO)_4$ photodissociation at 261 nm. *The Journal of Chemical Physics* **2022**, *156*, 144306.

(285) Whetten, R. L.; Fu, K. J.; Grant, E. R. Pulsed-laser photocatalytic isomerization and hydrogenation of olefins. *Journal of the American Chemical Society* **1982**, *104*, 4270–4272.

(286) Bitterwolf, T. E. Organometallic photochemistry at the end of its first century. *Journal of Organometallic Chemistry* **2004**, *689*, 3939–3952.

(287) Glascoe, E. A.; Sawyer, K. R.; Shanoski, J. E.; Harris, C. B. The Influence of the Metal Spin State in the Iron-Catalyzed Alkene Isomerization Reaction Studied with Ultrafast Infrared Spectroscopy. *The Journal of Physical Chemistry C* **2007**, *111*, 8789–8795.

(288) *Photophysics of organometallics*; Lees, A. J., Ed.; Springer Science & Business Media: Berlin and Heidelberg, Germany, 2010; Vol. 29.

(289) Nasri Lari, H. et al. Shedding light on iron pentacarbonyl photochemistry through a CVD case study. *Catalysis Communications* **2017**, *100*, 19–23.

(290) Whetten, R. L.; Fu, K.-J.; Grant, E. R. Photodissociation dynamics of $Fe(CO)_5$: Excited state lifetimes and energy disposal. *The Journal of Chemical Physics* **1983**, *79*, 4899–4911.

(291) Seder, T. A.; Ouderkirk, A. J.; Weitz, E. The wavelength dependence of excimer laser photolysis of $Fe(CO)_5$ in the gas phase. Transient infrared spectroscopy and kinetics of the $Fe(CO)_x$ (x = 4, 3, 2) photofragments. *The Journal of Chemical Physics* **1986**, *85*, 1977–1986.

(292) Poliakoff, M.; Weitz, E. Shedding light on organometallic reactions: the characterization of tetracarbonyliron $(Fe(CO)_4)$, a prototypical reaction intermediate. *Accounts of Chemical Research* **1987**, *20*, 408–414.

(293) Bañares, L.; Baumert, T.; Bergt, M.; Kiefer, B.; Gerber, G. The ultrafast photodissociation of Fe(CO)$_5$ in the gas phase. *The Journal of Chemical Physics* **1998**, *108*, 5799–5811.

(294) Leadbeater, N. Enlightening organometallic chemistry: the photochemistry of Fe(CO)$_5$ and the reaction chemistry of unsaturated iron carbonyl fragments. *Coordination Chemistry Reviews* **1999**, *188*, 35–70.

(295) Snee, P. T.; Payne, C. K.; Mebane, S. D.; Kotz, K. T.; Harris, C. B. Dynamics of Photosubstitution Reactions of Fe(CO)$_5$: An Ultrafast Infrared Study of High Spin Reactivity. *Journal of the American Chemical Society* **2001**, *123*, 6909–6915.

(296) Long, C. In *Photophysics of Organometallics*, Lees, A. J., Ed.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2010, pp 159–191.

(297) Fuss, W.; Trushin, S.; Schmid, W. Ultrafast photochemistry of metal carbonyls. *Research on Chemical Intermediates* **2001**, *27*, 447–457.

(298) Cole-Filipiak, N. C.; Schrader, J. T. P.; McCaslin, L. M.; Ramasesha, K. Ultraviolet photodissociation of gas-phase iron pentacarbonyl probed with ultrafast infrared spectroscopy. *Journal of Chemical Physics* **2021**, *154*, 134308.

(299) Daniel, C.; Bénard, M.; Dedieu, A.; Wiest, R.; Veillard, A. Theoretical aspects of the photochemistry of organometallics. 3. Potential energy curves for the photodissociation of Fe(CO)$_5$. *Journal of Physical Chemistry* **1984**, *88*, 4805–4811.

(300) González-Blanco, O.; Branchadell, V. Density functional study of the Fe–CO bond dissociation energies of Fe(CO)$_5$. *The Journal of Chemical Physics* **1999**, *110*, 778–783.

(301) Rubner, O.; Engel, V.; Hachey, M.; Daniel, C. A CASSCF/MR-CCI study of the excited states of Fe(CO)$_5$. *Chemical Physics Letters* **1999**, *302*, 489–494.

(302) Goumans, T. P. M.; Ehlers, A. W.; van Hemert, M. C.; Rosa, A.; Baerends, E.-J.; Lammertsma, K. Photodissociation of the Phosphine-Substituted Transition Metal-Carbonyl Complexes Cr(CO)$_5$ and Fe(CO)$_4$: A Theoretical Study. *Journal of the American Chemical Society* **2003**, *125*, 3558–3567.

(303) Persson, B. J.; Roos, B. O.; Pierloot, K. A theoretical study of the chemical bonding in M(CO)$_x$ (M = Cr, Fe, and Ni). *The Journal of Chemical Physics* **1994**, *101*, 6810–6821.

(304) Marquez, A.; Daniel, C.; Fernandez Sanz, J. The vacuum-ultraviolet spectrum of iron pentacarbonyl: an experimental analysis supported by a CASSCF CCI study of the Rydberg states. *The Journal of Physical Chemistry* **1992**, *96*, 121–123.

(305) Banerjee, A. et al. Photoinduced bond oscillations in ironpentacarbonyl give delayed synchronous bursts of carbonmonoxide release. *Nature Communications 2022 13:1* **2022**, *13*, 1–10.

(306) Ryther, R. J.; Weitz, E. Reaction kinetics of coordinatively unsaturated iron carbonyls formed on gas-phase excimer laser photolysis of iron pentacarbonyl. *The Journal of Physical Chemistry* **1991**, *95*, 9841–9852.

(307) Trushin, S. A.; Fuss, W.; Kompa, K. L.; Schmid, W. E. Femtosecond Dynamics of Fe(CO)$_5$ Photodissociation at 267 nm Studied by Transient Ionization. *The Journal of Physical Chemistry A* **2000**, *104*, 1997–2006.

(308) Ihee, H.; Cao, J.; Zewail, A. H. Ultrafast electron diffraction of transient Fe(CO)$_4$: Determination of molecular structure and reaction pathway. *Angewandte Chemie* **2001**, *113*, 1580–1584.

(309) Wernet, P. et al. Communication: Direct evidence for sequential dissociation of gas-phase Fe(CO)5 via a singlet pathway upon excitation at 266 nm. *Journal of Chemical Physics* **2017**, *146*, 211103.

(310) Leitner, T. et al. Time-resolved electron spectroscopy for chemical analysis of photodissociation: Photoelectron spectra of Fe(CO)$_5$, Fe(CO)$_4$, and Fe(CO)$_3$. *The Journal of Chemical Physics* **2018**, *149*, 044307.

(311) Herbert, J. M.; Zhang, X.; Morrison, A. F.; Liu, J. Beyond Time-Dependent Density Functional Theory Using only Single Excitations: Methods for Computational Studies of Excited States in Complex Systems. *Accounts of Chemical Research* **2016**, *49*, 931–941.

(312) Zhang, X.; Herbert, J. M. Spin-flip, tensor equation-of-motion configuration interaction with a density-functional correction: A spin-complete method for exploring excited-state potential energy surfaces. *Journal of Chemical Physics* **2015**, *143*, 234107.

(313) Levine, D. S.; Hait, D.; Tubman, N. M.; Lehtola, S.; Whaley, K. B.; Head-Gordon, M. CASSCF with Extremely Large Active Spaces Using the Adaptive Sampling Configuration Interaction Method. *Journal of Chemical Theory and Computation* **2020**, *16*, 2340–2354.

(314) Dohm, S.; Hansen, A.; Steinmetz, M.; Grimme, S.; Checinski, M. P. Comprehensive Thermochemical Benchmark Set of Realistic Closed-Shell Metal Organic Reactions. *Journal of Chemical Theory and Computation* **2018**, *14*, 2596–2608.

(315) Iron, M. A.; Janes, T. Evaluating Transition Metal Barrier Heights with the Latest Density Functional Theory Exchange–Correlation Functionals: The MOBH35 Benchmark Database. *The Journal of Physical Chemistry A* **2019**, *123*, 3761–3781.

(316) Chan, B.; Gill, P. M. W.; Kimura, M. Assessment of DFT Methods for Transition Metals with the TMC151 Compilation of Data Sets and Comparison with Accuracies for Main-Group Chemistry. *Journal of Chemical Theory and Computation* **2019**, *15*, 3610–3622.

(317) Carter-Fenk, K.; Shee, J.; Head-Gordon, M. Optimizing the regularization in size-consistent second-order Brillouin-Wigner perturbation theory. *The Journal of chemical physics* **2023**, *159*, 171104.

(318) Carter-Fenk, K.; Head-Gordon, M. Repartitioned Brillouin-Wigner perturbation theory with a size-consistent second-order correlation energy. *Journal of Chemical Physics* **2023**, *158*, 234108.

(319) Head-Gordon, M.; Rico, R. J.; Oumi, M.; Lee, T. J. A doubles correction to electronic excited states from configuration interaction in the space of single substitutions. *Chemical Physics Letters* **1994**, *219*, 21–29.

(320) Christiansen, O.; Koch, H.; Jørgensen, P. The second-order approximate coupled cluster singles and doubles model CC2. *Chemical Physics Letters* **1995**, *243*, 409–418.

(321) Stanton, J. F.; Gauss, J. Perturbative treatment of the similarity transformed Hamiltonian in equation-of-motion coupled-cluster approximations. *The Journal of Chemical Physics* **1995**, *103*, 1064–1076.

(322) Grell, G. et al. Multi-reference approach to the calculation of photoelectron spectra including spin-orbit coupling. *Journal of Chemical Physics* **2015**, *143*, 74104.

(323) Kasper, J. M.; Stetina, T. F.; Jenkins, A. J.; Li, X. Ab initio methods for L-edge x-ray absorption spectroscopy. *Chemical Physics Reviews* **2020**, *1*, 11304.

(324) Jenkins, A. J.; Hu, H.; Lu, L.; Frisch, M. J.; Li, X. Two-Component Multireference Restricted Active Space Configuration Interaction for the Computation of L-Edge X-ray Absorption Spectra. *Journal of Chemical Theory and Computation* **2022**, *18*, 141–150.

(325) Grofe, A.; Li, X. Relativistic nonorthogonal configuration interaction: application to L2,3-edge X-ray spectroscopy. *Physical Chemistry Chemical Physics* **2022**, *24*, 10745–10756.

(326) Vidal, M. L.; Coriani, S.; Pokhilko, P.; Krylov, A. I. Equation-of-motion coupled-cluster theory to model l-edge x-ray absorption and photoelectron spectra. *Journal of Physical Chemistry Letters* **2020**, *11*, 8314–8321.

(327) Dery Folkestad, S. et al. Enhanced understanding of X-ray absorption in liquid water using novel coupled cluster methodologies. *arXiv e-prints* **2023**, arXiv:2308.09495, arXiv:2308.09495.