Andreas Sæther Skeidsvoll

Real-time coupled-cluster approaches for electronic multiphoton processes in atoms and molecules

Doctoral thesis

NTNU
Norwegian University of
Science and Technology

Andreas Sæther Skeidsvoll

Real-time coupled-cluster approaches for electronic multiphoton processes in atoms and molecules

Thesis for the Degree of Philosophiae Doctor

Trondheim, July 2023

Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Chemistry

NTNU

Norwegian University of
Science and Technology

Abstract

This thesis details the implementation and application of the time-dependent coupled cluster (TDCC) method and two variants of the time-dependent equation-of-motion coupled-cluster (TD-EOM-CC) method in the recently released quantum chemistry program e^T. The methods are used to simulate interactions between atomic and molecular electrons and electromagnetic fields, on the attosecond-to-femtosecond timescale and for different field strengths, with the purpose of enhancing the knowledge of photochemistry and the mathematical characteristics of the different real-time coupled-cluster (RT-CC) methods.

In particular, we simulate transient absorption of attosecond x-ray pulses by molecular electrons following electronic excitation by ultrashort ultraviolet-visible pulses, with both TDCC and TD-EOM-CC. For the TD-EOM-CC simulations, the time-dependent state is expanded in a basis of field-free EOM-CC states obtained through an approach based on the asymmetric band Lanczos algorithm, which targets transitions involved in the interaction with the two pulses. For the TDCC simulations, the x-ray absorption is related to the phases of the valence-excited superposition, and analyzed in terms of the excitation energies of the field-free EOM-CC states. For the TD-EOM-CC simulations, the fluctuating electron density, approximated by a projected expectation value of the dipole moment vector, correlates with the time-dependence of the absorption integrated over relevant x-ray frequencies. The results demonstrate that real-time coupled cluster methods can be used for modeling attosecond spectroscopy in the weakly nonlinear regime, and support the idea that x-ray pulses can be used as local probes of the time-dependent electron density.

Furthermore, the TDCC and TD-EOM-CC methods are used for simulating the interaction of atomic and molecular electrons with strong fields. From simulations of collective Rabi oscillations, the TD-EOM-CC method is shown to be more stable than the TDCC method, for which certain time-dependent parameters blow up during the integration. The stability of TD-EOM-CC does however come at a cost, as the errors in the expectation values are seen to increase with the number of subsystems in resonance

with the field. However, subsystems out of resonance with the field do not impact the time-dependent energy expectation value, suggesting that TD-EOM-CC can be used for simulating single-subsystem transitions even in strong fields and for non-zero subsystem interactions. Moreover, through both theoretical analysis and simulations it is demonstrated that TDCC exhibits the correct scaling properties, even when dealing with multiple subsystems in resonance with the field.

The TD-EOM-CC method is also used to model process of stimulated x-ray Raman scattering (ISXRS) by atoms and molecules, where the TD-EOM-CC state is analyzed in terms of field-free EOM-CC state populations and visualizations of the time-dependent one-electron density in real space. The populations are shown to be sensitive to the polarization and carrier frequency of the external field, and the one-electron density displays a behavior agreeing with the expected behavior of a valence-excited wavepacket generated through ISXRS.

Acknowledgments

I would like to express my heartfelt gratitude to Prof. Henrik Koch for his excellent supervision of my PhD project. His creativity, enthusiasm and extensive knowledge have been instrumental in making this work possible. I would also like to thank Sarai D. Folkestad, Eirik F. Kjønstad, Rolf H. Myhre, and my co-supervisor Ida-Marie Høyvik for their invaluable support and mentorship during my early days in the Quantum Chemistry group at NTNU and with the e^T development.

I am sincerely grateful to my collaborators, including Alice Balbi, Torsha Moitra, Sonia Coriani, and Alexander C. Paul, and all the other authors of the e^T program, for their invaluable contributions in our collaborative efforts. Special thanks to Alexander C. Paul for his assistance in proofreading an earlier version of this thesis. I also wish to extend my gratitude to my colleagues at the Department of Chemistry at NTNU and at Scuola Normale Superiore in Italy for the stimulating scientific discussions we shared and the memorable experiences we had together.

Furthermore, I want to express my profound appreciation to my beloved wife Solveig, whose unwavering love and patience have been an invaluable pillar of strength throughout this journey. Raising our little son Markus together has brought me tremendous joy, and I cherish every moment spent with my family. I am also grateful to my parents, siblings, in-laws, and friends for their support and encouragement throughout the years. Your presence in my life is of immeasurable importance to me, and I am truly blessed to have you all by my side.

Lastly, I would like to thank the Department of Chemistry at NTNU for providing me with the opportunity to pursue my PhD studies, and the defense committee for dedicating their time to evaluate this thesis. I hope that you, as well as all other readers, will find it informative and engaging.

Contents

Chapter 1

Introduction

The true dynamical behavior of molecules is hard to unravel. Not only because of their small size and short timescale of motion, but also due to the diverse ways in which a molecule can interact, both internally and with its environment. For some processes occurring in nature, the molecular behavior can be well-represented by simple models, making use of empirical parameters and classical equations of motion [1–3]. Other molecular processes, involving for instance charge migration and bond breaking, can demand a more precise description.

The most precise description of molecular behavior is given by quantum theory. Unfortunately, the computational resources needed for full quantum simulation of large molecules quickly grow beyond the reach of present-day computers [4]. However, quantum methods can in many cases be approximated through the identification and omission of less important degrees of freedom. A primary goal of quantum chemistry to develop practical approximate methods that can accurately predict and explain the behavior of molecules in a broad range of chemical processes. Together with careful experimentation and analysis, such models can hopefully increase the conceptual understanding of chemistry.

One of the frontiers of quantum chemistry is the proper description of molecules exposed to short and strong laser pulses. Such pulses are often used in experiments that aim to give an increased understanding of photochemistry, since they can quickly probe and change the state of a molecule. The pulses can for instance be used to make the analog of movies of molecular motion [5], by combining measurements by pulses at different time delays. Strong pulses can also be used to quickly control the molecular state, which can lead to new ways of doing chemistry [6]. The simulation of molecules interacting with short and strong laser pulses is often done in the time domain, since ultrafast and multiphotonic processes can be studied in a non-perturbative manner, and time-dependent properties can be

calculated straightforwardly.

Free-electron lasers and tabletop high-harmonic generation sources can now produce x-ray pulses with durations in the range of attoseconds, enabled by the short periods of x-ray radiation [7, 8]. Core excitations resulting from molecular x-ray absorption are generally sensitive to the density around a particular nucleus [9]. As a result, pulsed x-ray spectroscopy has the potential to accurately resolve molecular charge migration, both temporally and spatially [10]. This motivates the further development and application of quantum chemical methods for simulating pulsed x-ray spectroscopy.

Traditional coupled-cluster methods originated in nuclear physics, and have since been both adopted and further popularized by the electronic structure community [11]. Untruncated coupled-cluster methods can describe the correlated system exactly in the given basis set. Truncated coupled-cluster methods, on the other hand, are approximate methods that often give accurate descriptions of correlated electronic structure [12], with properties and computational costs that scale reasonably with respect to the size of the quantum system.

Coupled-cluster methods have traditionally been treated in the frequency domain, and the molecular response to external interactions calculated perturbatively. However, the number of studies which treat coupled-cluster methods in the real-time domain has grown recently [13], including approaches based on the time-dependent coupled-cluster (TDCC) method [14] and the time-dependent equation-of-motion coupled-cluster (TD-EOM-CC) method [15, 16]. In this thesis, truncated TDCC and TD-EOM-CC methods have been implemented and used for simulating molecular electrons interacting with laser pulses. The methods have been compared, revealing that truncated TD-EOM-CC methods can fail to describe strong-field interactions, in a distinct manner from the known numerical instability issues of TDCC methods [17, 18].

In Chapter 2 the theory of quantum mechanics is introduced, and its use in describing atoms and molecules is explained. A classical description of the electromagnetic field and its semiclassical interaction with atomic and molecular electrons in the length gauge is given in Chapter 3. Chapter 4 provides an introduction to the solution of the electronic time-independent Schrödinger equation for atoms and molecules. In Chapter 5, time-independent and time-dependent coupled-cluster methods are introduced, which are used for the simulations in this thesis. Chapter 6 introduces select topics in semiclassical laser-molecule interactions that serve a basis for understanding the physical effects observed in our simulations. The findings of the papers in this thesis are summarized and

concluded in Chapter 7.

Chapter 2

Quantum mechanical description of molecules

In this chapter, an introduction is given on the description of atoms and molecules in the framework of traditional quantum mechanics, with an emphasis on the time-dependence of the atomic and molecular electrons.

2.1 Vectors and operators in Dirac notation

The pure quantum state of a N-particle system can, in the Schrödinger picture and Dirac notation, be represented by the time-dependent "ket"

$$|\Psi(t)\rangle \tag{2.1}$$

residing in a Hilbert space [19], and its dual, the "bra"

$$\langle\Psi(t)| = |\Psi(t)\rangle^{\dagger}. \tag{2.2}$$

The dagger is used to denote Hermitian conjugation, which is the combined operation of complex conjugation and transposition.

Observable properties of the system are represented by linear Hermitian operators

$$O(t) = O^{\dagger}(t), \tag{2.3}$$

which can exhibit an explicit dependence on time. The bra and ket are used to form inner products in the Hilbert space, possibly involving operators. In accordance with the Born rule, physical measurements have their theoretical analog in the inner product

$$\langle O\rangle(t) = \langle\Psi(t)|O(t)|\Psi(t)\rangle, \tag{2.4}$$

which gives the sum of the probabilities of the different outcomes of the measurement, which can vary with time. The probabilities can be compared to repeated measurements made in experiments by an observer. The order of appearance of operators in equations matters, as operators do not necessarily commute, $O_1(t)O_2(t) \neq O_2(t)O_1(t)$. The lacking commutativity of certain operators sets fundamental limits on how precise simultaneous measurements of the corresponding observables can be made. This forms the foundation for the different variants of the Heisenberg uncertainty principle. For more information about bras, kets and operators in quantum mechanics, see Ref. [19]

2.2   Time-dependent Schrödinger equation

The time evolution of the non-relativistic N-particle ket $|\Psi(t)\rangle$ is governed by the time-dependent Schrödinger equation (TDSE) [20]

$$i\frac{\mathrm{d}}{\mathrm{d}t}|\Psi(t)\rangle = H(t)|\Psi(t)\rangle, \tag{2.5}$$

where the linear Hermitian operator $H(t)$ is the Hamiltonian, the generator of the time evolution of the system. Assuming that the state of the system is known at $t = t_0$, the time-dependent state at time t can be written as the time evolution operator $U(t, t_0)$ acting on the initial state $|\Psi(t_0)\rangle$,

$$|\Psi(t)\rangle = U(t, t_0)|\Psi(t_0)\rangle. \tag{2.6}$$

Inserting Equation (2.6) into Equation (2.5) before projecting onto $\langle\Psi(t_0)|$, we get the following equation for the time evolution operator

$$i\frac{\mathrm{d}}{\mathrm{d}t}U(t, t_0) = H(t)U(t, t_0). \tag{2.7}$$

Formally, the solution to this equation can be written as

$$U(t, t_0) = \mathcal{T}e^{-i\int_{t_0}^{t}\mathrm{d}t_1 H(t_1)}, \tag{2.8}$$

where the time-ordering procedure, denoted by $\mathcal{T}$, orders all operators that correspond to earlier time values to the right of the later ones. If the Hamiltonian is time-independent, the time-evolution operator has the simpler form

$$U(t, t_0) = e^{-iH(t-t_0)}, \tag{2.9}$$

leading to the time-dependent state

$$|\Psi(t)\rangle = e^{-iH(t-t_0)}|\Psi(t_0)\rangle. \tag{2.10}$$

For more details about the TDSE, see Chapter 3 of Ref. [20].

2.3 Position-spin representation

Until now, we have described the quantum states and operators in terms of abstract quantities, without reference to their explicit representation in terms of mathematical functions. Such a representation is needed in order to perform numerical simulations of quantum systems.

A useful representation of quantum states of systems with N distinguishable particles can be found through resolution of the identity operator in terms of the quantum eigenstates belonging to the position-spin coordinates $x_j = (r_j, \sigma_j)$ of the N particles [20], where r_j is the position and σ_j the spin of particle j,

$$\mathbb{1} = \int \mathrm{d}x_1 \cdots \mathrm{d}x_N \, |x_1, \ldots, x_N\rangle \, \langle x_1, \ldots, x_N| . \qquad (2.11)$$

By application of the identity operator onto the ket and bra, we get the following identities

$$|\Psi(t)\rangle = \int \mathrm{d}x_1 \cdots \mathrm{d}x_N \, |x_1, \ldots, x_N\rangle \, \Psi(x_1, \ldots, x_N, t), \qquad (2.12)$$

$$\langle \Psi(t)| = \int \mathrm{d}x_1 \cdots \mathrm{d}x_N \, \Psi^*(x_1, \ldots, x_N, t) \, \langle x_1, \ldots, x_N| , \qquad (2.13)$$

where the function

$$\Psi(x_1, \ldots, x_N, t) = \langle x_1, \ldots, x_N | \Psi(t)\rangle , \qquad (2.14)$$

is called the wavefunction of the quantum state in the position-spin representation, and the asterisk denotes complex conjugation. The quantity

$$\rho(x_1, \ldots, x_N, t) = |\Psi(x_1, \ldots, x_N, t)|^2 \qquad (2.15)$$

is the probability density associated with observing the particles with indices $1, \ldots, N$ at the respective coordinates $x_1, \ldots, x_N$, at time t. For identical particles, Equation (2.15) corresponds to the probability density associated with observing one of the identical particles at x_1, another one at x_2, and so on. The additional normalization factor $1/N!$ is then needed on the right-hand side of the resolution of identity in Equation (2.11) [20].

In the position-spin representation, the TDSE takes the following form

$$i\frac{\mathrm{d}}{\mathrm{d}t}\Psi(x_1, \ldots, x_N, t) = H(x_1, \ldots, x_N, t)\Psi(x_1, \ldots, x_N, t). \qquad (2.16)$$

2.4   Minimal coupling molecular Hamiltonian

Assuming no external interactions, the non-relativistic and spin-free Hamiltonian of an atom or a molecule can in the position-spin representation be written as

$$H(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N,t) = -\sum_i \frac{1}{2m_i}\nabla_i^2 + \frac{1}{2}\sum_{i\neq j}\frac{q_i q_j}{|\boldsymbol{r}_i - \boldsymbol{r}_j|}, \qquad (2.17)$$

where m_i is the mass, q_i the charge, $\boldsymbol{r}_i$ the position and $\boldsymbol{\nabla}_i = \frac{\partial}{\partial x_i} + \frac{\partial}{\partial y_i} + \frac{\partial}{\partial z_i}$ the gradient along the Cartesian coordinates of particle i. Note that this Hamiltonian is time-independent.

The minimal coupling scheme can be used to introduce the interaction with a time-dependent electromagnetic field into the Hamiltonian. As an introduction to this scheme, we assess the effect of performing the gauge transformation

$$\Psi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N,t) \rightarrow \Psi'(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N,t) = e^{i\sum_i q_i\chi(\boldsymbol{x}_i,t)}\Psi(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N,t), \ (2.18)$$

in the TDSE of the molecular Hamiltonian, where $\chi(\boldsymbol{r},t)$ is a local and time-dependent scalar gauge field. By pre-multiplying the TDSE by $e^{i\sum_i q_i\chi(\boldsymbol{r}_i,t)}$, we can identify the transformed Hamiltonian H' as [21]

$$\begin{aligned}
H'&(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N,t)\\
&= e^{i\sum_i q_i\chi(\boldsymbol{r}_i,t)}H(\boldsymbol{x},t)e^{-i\sum_i q_i\chi(\boldsymbol{r}_i,t)} - \sum_i q_i\frac{\mathrm{d}\chi(\boldsymbol{r}_i,t)}{\mathrm{d}t}\\
&= \sum_i \frac{1}{2m_i}\big(\boldsymbol{p}_i - q_i\boldsymbol{\nabla}\chi(\boldsymbol{r}_i,t)\big)^2 + \frac{1}{2}\sum_{i\neq j}\frac{q_i q_j}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} - \sum_i q_i\frac{\mathrm{d}\chi(\boldsymbol{r}_i,t)}{\mathrm{d}t},
\end{aligned} \qquad (2.19)$$

which implies that the Hamiltonian is not invariant to this gauge transformation.

Introducing the electromagnetic field into the Hamiltonian can restore the gauge invariance. In classical electromagnetic theory, the physically observable electric $\boldsymbol{E}(\boldsymbol{r},t)$ and magnetic $\boldsymbol{B}(\boldsymbol{r},t)$ fields can be related to the vector $\boldsymbol{A}(\boldsymbol{r},t)$ and scalar $\varphi(\boldsymbol{r},t)$ potentials through the equations

$$\boldsymbol{E}(\boldsymbol{r},t) = -\boldsymbol{\nabla}\varphi(\boldsymbol{r},t) - \frac{\partial \boldsymbol{A}(\boldsymbol{r},t)}{\partial t}, \qquad (2.20)$$

$$\boldsymbol{B}(\boldsymbol{r},t) = \boldsymbol{\nabla} \times \boldsymbol{A}(\boldsymbol{r},t), \qquad (2.21)$$

which are invariant to the following gauge transformations

$$\varphi(\boldsymbol{r},t) \rightarrow \varphi'(\boldsymbol{r},t) = \varphi(\boldsymbol{r},t) - \frac{\mathrm{d}\chi(\boldsymbol{r},t)}{\mathrm{d}t}, \qquad (2.22)$$

$$\boldsymbol{A}(\boldsymbol{r},t) \rightarrow \boldsymbol{A}'(\boldsymbol{r},t) = \boldsymbol{A}(\boldsymbol{r},t) + \boldsymbol{\nabla}\chi(\boldsymbol{r},t). \qquad (2.23)$$

Taking the gauge transformation properties of the vector and scalar potentials into account, one can argue that the following form of the Hamiltonian is reasonable

$$H(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N,t) = \sum_i \frac{1}{2m_i}\left(\boldsymbol{p}_i + q_i \boldsymbol{A}(\boldsymbol{r}_i,t)\right)^2 + \sum_{i>j}\frac{q_i q_j}{\left|\boldsymbol{r}_i - \boldsymbol{r}_j\right|} + \sum_i q_i \varphi(\boldsymbol{r}_i,t),$$
(2.24)

since Equation (2.18), Equation (2.23) and Equation (2.22) imply that the Hamiltonian is invariant under gauge transformations,

$$H(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N,t) \to H'(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N,t) = H(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N,t).$$
(2.25)

The coupling of the classical electromagnetic field to the molecule given in Equation (2.24) is known as minimal coupling, since it does not involve higher multipole moments of the charge distribution.

2.5 Clamped nuclei

The coordinates of the molecule can be separated into the electronic coordinates $\boldsymbol{x},\ldots,\boldsymbol{x}_{N_e}$ and the nuclear coordinates $\boldsymbol{x}_{N_e+1},\ldots,\boldsymbol{x}_{N_e+N_n}$. For ease of notation, we denote the electronic coordinates by $\boldsymbol{x}$ and the nuclear coordinates by $\boldsymbol{X}$. The molecular Hamiltonian Equation (2.17) can be written as

$$H(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N) = H_0(\boldsymbol{x},\boldsymbol{X}) + T_n(\boldsymbol{X}),$$
(2.26)

where the term corresponding to the nuclear kinetic energy has been separated out,

$$T_n(\boldsymbol{X}) = -\sum_I \frac{1}{2M_I}\nabla_I^2,$$
(2.27)

with the index I running over the nuclear coordinates only. The other term is known as the electronic Hamiltonian,

$$H_0(\boldsymbol{x},\boldsymbol{X}) = -\frac{1}{2}\sum_i \nabla_i^2 + \frac{1}{2}\sum_{i\neq j}\frac{1}{\left|\boldsymbol{r}_i - \boldsymbol{r}_j\right|} - \sum_{iJ}\frac{Z_J}{\left|\boldsymbol{r}_i - \boldsymbol{R}_J\right|} + \frac{1}{2}\sum_{I\neq J}\frac{Z_I Z_J}{\left|\boldsymbol{R}_I - \boldsymbol{R}_J\right|},$$
(2.28)

and includes the kinetic energy of the electrons, indexed by i, and the Coulomb interaction between all particle pairs. It only depends parametrically on the nuclear coordinates.

The atomic and molecular Hilbert space typically exhibits high dimensionality, with the kinetic energy of the nuclei often significantly smaller

than the total atomic or molecular energy [22]. As a result, we can simplify the system by invoking the clamped nuclei approximation, which involves neglecting the nuclear kinetic energy term. This approximation implies that

$$H(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N) \approx H_0(\boldsymbol{x},\boldsymbol{X}), \qquad (2.29)$$

thus reducing the description of the system to the electronic Hilbert space.

The Born-Oppenheimer approximation involves the clamped nuclei approximation as its first step, allowing for the description of nuclear dynamics under the assumption that electrons adjust adiabatically to the moving nuclei. Although the original formulation of the Born-Oppenheimer approximation does not account for time-dependent interactions, formulations including such interactions have been proposed [23]. For simplicity, however, we choose to avoid the description of the nuclear dynamics. By applying the minimal coupling scheme to the electronic Hamiltonian in Equation (2.28) and invoking the clamped nuclei approximation, the total atomic or molecular Hamiltonian can be expressed as

$$
\begin{aligned}
H(\boldsymbol{x},\boldsymbol{X},t) \\
\approx -\frac{1}{2}\sum_i \left(\nabla_i^2 + q_i \boldsymbol{A}(\boldsymbol{r}_i,t)\right) + \frac{1}{2}\sum_{i\neq j}\frac{1}{|\boldsymbol{r}_i-\boldsymbol{r}_j|} - \sum_{iJ}\frac{Z_J}{|\boldsymbol{r}_i-\boldsymbol{R}_J|} \\
+ \frac{1}{2}\sum_{I\neq J}\frac{Z_I Z_J}{|\boldsymbol{R}_I-\boldsymbol{R}_J|} + \sum_i q_i \varphi(\boldsymbol{r}_i,t).
\end{aligned} \qquad (2.30)
$$

Chapter 3

Semiclassical description of laser-molecule interactions

In semiclassical physics, the system is treated in two parts, where one part is described quantum mechanically and the other one classically. We choose to describe the atomic and molecular electrons quantum mechanically, and the external electromagnetic field as a classical field described by Maxwell's equations. In the following, the mathematical description of the external field and its semiclassical interaction with the electrons is introduced.

3.1 Wave equation for the vector potential

The homogeneous Maxwell's equations are given by [24]

$$\boldsymbol{\nabla} \cdot \boldsymbol{B}(\boldsymbol{r}, t) = 0, \tag{3.1}$$

$$\boldsymbol{\nabla} \times \boldsymbol{E}(\boldsymbol{r}, t) + \frac{\partial \boldsymbol{B}(\boldsymbol{r}, t)}{\partial t} = 0. \tag{3.2}$$

By construction, the vector $\boldsymbol{A}(\boldsymbol{r}, t)$ and scalar $\varphi(\boldsymbol{r}, t)$ potentials satisfy these equations, which can be seen by insertion of Equation (2.20) and Equation (2.21). In the following, we derive additional conditions for the scalar and vector potentials. Assuming that the electromagnetic field is itself without sources, that is, that the charge density $\rho(\boldsymbol{r}, t)$ and the current density $\boldsymbol{J}(\boldsymbol{r}, t)$ are zero, the inhomogeneous Maxwell's equations take the following form [24]

$$\boldsymbol{\nabla} \cdot \boldsymbol{E}(\boldsymbol{r}, t) = 0, \tag{3.3}$$

$$\boldsymbol{\nabla} \times \boldsymbol{B}(\boldsymbol{r}, t) - \frac{1}{c^2} \frac{\partial \boldsymbol{E}(\boldsymbol{r}, t)}{\partial t} = 0, \tag{3.4}$$

where $c = 1/\sqrt{\mu_0 \epsilon_0}$ is the speed of light in vacuum, μ_0 the vacuum permeability and ϵ_0 the vacuum permittivity. We choose to partially fix the gauge of the scalar and vector potentials through the Coulomb gauge condition

$$\nabla \cdot A(r, t) = 0. \tag{3.5}$$

By thereafter inserting Equation (2.20) and Equation (3.5) into Equation (3.3), we obtain Laplace's equation for the scalar potential,

$$\nabla^2 \varphi(r, t) = 0. \tag{3.6}$$

By imposing the boundary condition of the scalar field $\varphi(r, t)$ vanishing at infinity, the gauge is completely fixed and Equation (3.6) has the unique solution

$$\varphi(r, t) = 0. \tag{3.7}$$

Inserting Equation (2.20), Equation (2.21) and Equation (3.7) into Equation (3.4), we obtain the homogeneous wave equation for the vector potential

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) A(r, t) = 0. \tag{3.8}$$

3.2   Plane waves

The vector potential $A(r, t)$ can be written as a linear combination of plane waves, which are arbitrary functions of $k \cdot r - \omega t$, where k is the wave vector and ω the angular frequency of the plane wave. In vacuum, where the vacuum dispersion relation $\omega = c|k|$ holds, the plane waves are solutions to Equation (3.8) and propagate without dispersion along the wave vector k [25]. The monochromatic plane waves are plane waves that can be written as [24]

$$\widetilde{A}_{k,\omega} e^{i(k \cdot r - \omega t)}, \tag{3.9}$$

where the coefficient $\widetilde{A}_{k,\omega}$ is a vector. The monochromatic plane waves form a complete basis for the vector potential $A(r, t)$, and in the limit of continuous summation, the coefficients $\widetilde{A}_{k,\omega}$ can be represented by the function $\widetilde{A}(k, \omega)$, and the general vector potential expressed as

$$A(r, t) = \int dk \int d\omega \, \widetilde{A}(k, \omega) e^{i(k \cdot r - \omega t)}. \tag{3.10}$$

3.3 Dipole approximation

The dot product $\boldsymbol{k} \cdot \boldsymbol{r}$ can be rewritten as

$$\boldsymbol{k} \cdot \boldsymbol{r} = kr(\hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{r}}), \tag{3.11}$$

where the magnitudes $k = |\boldsymbol{k}|$ and $r = |\boldsymbol{r}|$, and the unit vectors $\hat{\boldsymbol{k}} = \boldsymbol{k}/k$ and $\hat{\boldsymbol{r}} = \boldsymbol{r}/r$. The magnitude of the wave vector k is known as the wavenumber, and can be related to the wavelength λ_k through $k = 2\pi/\lambda_k$, giving

$$\boldsymbol{k} \cdot \boldsymbol{r} = \frac{2\pi r}{\lambda_k}(\hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{r}}), \tag{3.12}$$

We assume that the electromagnetic radiation of a given wavelength only interacts in a spatial region that is much smaller than the corresponding wavelength, $2\pi r/\lambda_k \ll 1$. Under this assumption, the use of the dipole approximation

$$\boldsymbol{k} \cdot \boldsymbol{r} \approx 0 \tag{3.13}$$

can be justified [26].

Inserting Equation (3.13) into Equation (3.10), the vector potential is approximated as

$$\boldsymbol{A}(\boldsymbol{r}, t) \approx \boldsymbol{A}(t) = \int \mathrm{d}\omega \, \widetilde{\boldsymbol{A}}(\omega) e^{-i\omega t}, \tag{3.14}$$

where we have defined $\widetilde{\boldsymbol{A}}(\omega) = \int \mathrm{d}\boldsymbol{k} \, \widetilde{\boldsymbol{A}}(\boldsymbol{k}, \omega)$. We can see that the electric field takes the shape of an electric dipole by inserting Equation (3.14) into Equation (2.20)

$$\boldsymbol{E}(\boldsymbol{r}, t) \approx \boldsymbol{E}(t) = \frac{\partial}{\partial t} \boldsymbol{A}(t) = \int \mathrm{d}\omega \, \widetilde{\boldsymbol{E}}(\omega) e^{-i\omega t} \tag{3.15}$$

where we have defined the function $\widetilde{\boldsymbol{E}}(\omega) = -i\omega \widetilde{\boldsymbol{A}}(\omega)$. From the freedom in choosing the function $\widetilde{\boldsymbol{A}}(\boldsymbol{k}, \omega)$, we can also choose $\widetilde{\boldsymbol{E}}(\omega)$ as an arbitrary function of ω. Inserting Equation (3.14) into Equation (2.21), we can see that the magnetic field disappears in the dipole approximation,

$$\boldsymbol{B}(\boldsymbol{r}, t) \approx \boldsymbol{\nabla} \times \boldsymbol{A}(t) = 0. \tag{3.16}$$

3.4   Carrier and envelope separation

In the dipole approximation, the time-dependent electric field $E(t)$ can be expressed with respect to the carrier frequency ω_0 by choosing the arbitrary function $\widetilde{E}(\omega)$ in Equation (3.15) as

$$\widetilde{E}(\omega) = \int \mathrm{d}t' \, \boldsymbol{F}_0(t') e^{i(\omega-\omega_0)t'}, \tag{3.17}$$

where $\boldsymbol{F}_0(t)$ is an arbitrary function of t. This choice gives the following form of the electric field

$$\begin{aligned}
\boldsymbol{E}(t) &= \int \mathrm{d}\omega \, \widetilde{E}(\omega) e^{-i\omega t} \\
&= \int \mathrm{d}t' \, \boldsymbol{F}_0(t') e^{-i\omega_0 t'} \int \mathrm{d}\omega \, e^{i\omega(t'-t)} \\
&= \boldsymbol{F}_0(t) e^{-i\omega_0 t},
\end{aligned} \tag{3.18}$$

where the time-dependent amplitude of the field oscillates as a monochromatic carrier wave at ω_0, modulated by the complex amplitude of $\boldsymbol{F}_0(t)$.

Furthermore, we assume that the field is linearly polarized,

$$\boldsymbol{F}_0(t) = F_0(t)\hat{\boldsymbol{\epsilon}}, \tag{3.19}$$

where $\hat{\boldsymbol{\epsilon}}$ is the real unit vector and $F_0(t)$ the complex amplitude in the polarization direction. We can write $F_0(t)$ in the complex form

$$F_0(t) = |F_0(t)| e^{-i\varphi(t)} \tag{3.20}$$

where the real and non-negative envelope function $|F_0(t)|$ is responsible for any amplitude modulation, and the real phase $\varphi(t) = -\arg(F_0)$ for any angle modulation. We furthermore assume that $\varphi(t)$ describes a constant offset between the phases of the carrier and envelope functions, $\varphi(t) = \varphi_{\mathrm{CEP}}$. Since the classical electric field is an observable, and thus real, we then obtain

$$\begin{aligned}
\boldsymbol{E}(t) &= \mathrm{Re}\left(|F_0(t)| e^{-i\omega t - i\varphi_{\mathrm{CEP}}}\right)\hat{\boldsymbol{\epsilon}} \\
&= |F_0(t)| \cos(\omega t + \varphi_{\mathrm{CEP}})\hat{\boldsymbol{\epsilon}}.
\end{aligned} \tag{3.21}$$

This equation still leaves the flexibility of choosing the envelope $|F_0(t)|$ of the electric field as any real and non-negative function of t, and the carrier-envelope phase φ_{CEP} as any real constant.
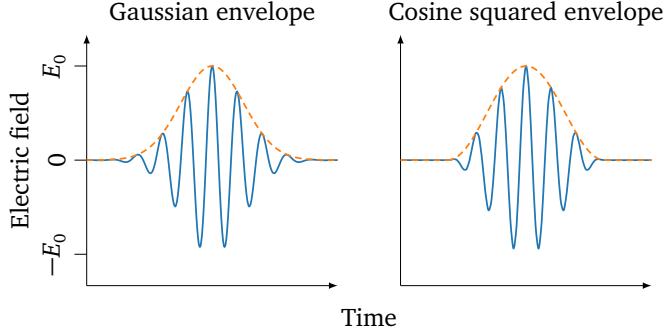
Figure 3.1: Illustration of pulses with Gaussian and cosine squared envelopes, and $\varphi_{\mathrm{CEP}} = 0$. The electric fields are shown in solid blue lines and the envelope functions in dashed orange lines.

3.5 Laser pulses

We model laser pulses with Equation (3.21), by choosing the envelope $|F_0(t)|$ as a function that is centered around a value of t, which we for simplicity take to be zero. One such choice of envelope is the Gaussian envelope

$$|F_0(t)| = E_0 e^{-t^2/(2\sigma^2)} \tag{3.22}$$

which is non-zero for all t. Another choice is the cosine squared envelope

$$|F_0(t)| = \begin{cases} E_0 \cos^2\left(\frac{\pi t}{T}\right), & -\frac{T}{2} \le t \le \frac{T}{2} \\ 0, & \text{otherwise,} \end{cases} \tag{3.23}$$

which is only non-zero within a single period of the squared cosine function. In Figure 3.1, we illustrate the magnitude $|E(t)|$ of the electric field for both Gaussian and cosine-squared envelopes, assuming $\varphi_{\mathrm{CEP}} = 0$.

In preparation for a discussion on the energy content of laser pulses, we note that the energy U contained in the electromagnetic field can be found by integrating the electromagnetic energy density $u(\mathbf{r}, t)$ over all space and time,

$$U = \int \mathrm{d}\mathbf{r} \int \mathrm{d}t \, u(\mathbf{r}, t). \tag{3.24}$$

In vacuum, the electromagnetic energy density is given by [24]

$$u(\mathbf{r}, t) = \frac{1}{2}\left(\epsilon_0 |\mathbf{E}(\mathbf{r}, t)|^2 + \frac{1}{\mu_0} |\mathbf{B}(\mathbf{r}, t)|^2 \right), \tag{3.25}$$

where ϵ_0 is the vacuum permittivity and μ_0 the vacuum permeability.

Within the dipole approximation, the electric field is assumed to be spatially uniform, and the magnetic field can be neglected. This approximation is generally valid when the size of the system interacting with a laser pulse is much smaller than the dominant wavelengths in the pulse. The energy density of the laser pulse can then be expressed as

$$u(\boldsymbol{r}, t) \approx \frac{\epsilon_0}{2}|\boldsymbol{E}(t)|^2 = \frac{\epsilon_0}{2}|F_0(t)|^2 \cos^2(\omega t + \varphi_{\mathrm{CEP}}), \qquad (3.26)$$

where the last equality is obtained by inserting Equation (3.21), which relates the electric field to the envelope and carrier wave of the pulse. However, the total energy derived from integrating this density over all space and time is infinite, since the integral is proportional to the integration volume. The integration can be limited to a specific volume that represents the spatial extent of the pulse, but the choice of this volume can be ambiguous because the dipole approximation is not valid for the entire volume of pulses with finite spatial extent. Instead of considering the total energy in the dipole approximation, it is more appropriate to focus on the energy density or the energy transfer to the system. To estimate the energy transfer, one can analyze the response of the system to the laser pulse by solving the time-dependent Schrödinger equation with the inclusion of the electric field interaction.

3.6   Length gauge

In the dipole approximation, the length gauge gives a simple description of semiclassical light-matter interaction, in which the classical electromagnetic field is represented only by the electric field $\boldsymbol{E}(t)$. The transformation from the Coulomb gauge to the length gauge can be performed with the gauge transformation $\chi(\boldsymbol{r}, t) = -\boldsymbol{r} \cdot \boldsymbol{A}(t)$, which inserted into Equation (2.22) and Equation (2.23) gives

$$\begin{aligned} \boldsymbol{A}'(\boldsymbol{r}, t) &= \boldsymbol{A}(t) + \nabla\big(-\boldsymbol{r} \cdot \boldsymbol{A}(t)\big) \\ &= 0, \end{aligned} \qquad (3.27)$$

$$\begin{aligned} \varphi'(\boldsymbol{r}, t) &= \varphi(\boldsymbol{r}, t) - \frac{\mathrm{d}}{\mathrm{d}t}\big(-\boldsymbol{r} \cdot \boldsymbol{A}(t)\big) \\ &= -\boldsymbol{r} \cdot \boldsymbol{E}(t), \end{aligned} \qquad (3.28)$$

where Equation (3.7) and Equation (3.15) are also used in obtaining the last equation.

Inserting these equations into the clamped nuclei Hamiltonian in Equation (2.30), we get the clamped nuclei length gauge Hamiltonian

$$H(\boldsymbol{x}, t) = H_0(\boldsymbol{x}) + V(\boldsymbol{x}, t), \tag{3.29}$$

where $H_0(\boldsymbol{x})$ is the field-free electronic Hamiltonian and

$$V(\boldsymbol{x}, t) = \sum_i \boldsymbol{r}_i \cdot \boldsymbol{E}(t) \tag{3.30}$$

describes the interaction of the electrons with the classical electromagnetic field.

3.7 Field-free eigenbasis

We can use the eigenstates of the field-free Hamiltonian H_0, represented by $|\psi_i\rangle$ and $\langle\psi_i|$, as a basis to express operators as matrices, with elements

$$O_{ij}(t) = \langle\psi_i| O(t) |\psi_j\rangle. \tag{3.31}$$

The time-dependent Hamiltonian $H(t)$ describing the electrons and their interaction with the field can also be written in this basis, with elements

$$H_{ij}(t) = \delta_{ij} E_j + V_{ij}(t), \tag{3.32}$$

where E_j is the eigenvalue of H_0 for $|\psi_i\rangle$ and the interaction term

$$V_{ij}(t) = -\boldsymbol{d}_{ij} \cdot \boldsymbol{E}(t), \tag{3.33}$$

where $\boldsymbol{d}_{ij}$ are the matrix elements of the electric dipole moment operator $\boldsymbol{d} = -\sum_i \boldsymbol{r}_i$.

Chapter 4

Solving the electronic time-independent Schrödinger equation

In real-time methods, the TDSE is solved by propagating the state of the quantum system from an initial state, which can often be constructed from the eigenstates of the field-free Hamiltonian. The temporal evolution of the time-dependent state can also often be expressed using these same eigenstates, as alluded to in Section 3.7. This chapter presents techniques for approximating these eigenstates, with a particular emphasis on the lowest-energy eigenstate corresponding to the ground state of the system. The discussion leads to the next chapter's topic of coupled cluster methods.

4.1 Time-independent Schrödinger equation

If the Hamiltonian in the position-spin representation is time-independent, $H(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N,t) = H(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)$, as is the case for the field-free electronic Hamiltonian, it is possible to separate the dependence on the coordinates $\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N$ and the time t. Following the method of separation of variables, we assume that the nth solution of the TDSE can be factorized into a product of a position-spin-dependent and a time-dependent function,

$$\Psi_n(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N,t) = \psi_n(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)\phi_n(t). \qquad (4.1)$$

Inserting this factorized solution into Equation (2.16), we obtain

$$\frac{H(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)\psi_n(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)}{\psi_n(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)} = \frac{\frac{\mathrm{d}\phi_n(t)}{\mathrm{d}t}}{\phi_n(t)} = E_n. \qquad (4.2)$$

The constant eigenvalue E_n is interpreted as the energy of the system in eigenstate n. From Equation (4.2), we get the time-dependent equation

$$\frac{\mathrm{d}\phi_n(t)}{\mathrm{d}t} = E_n\phi_n(t), \tag{4.3}$$

which, assuming that $\phi_n(t_0) = 1$, has the phase factor solution

$$\phi_n(t) = e^{-iE_n(t-t_0)}. \tag{4.4}$$

From Equation (4.2), we also get the stationary eigenvalue equation known as the time-indepedent Schrödinger equation (TISE)

$$H(\boldsymbol{x})\psi_n(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N) = E_n\psi_n(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N). \tag{4.5}$$

Since the Hamiltonian is Hermitian, it can according to the spectral theorem be diagonalized, and the solutions can be made orthogonal

$$\int \mathrm{d}\boldsymbol{x}_1\cdots\mathrm{d}\boldsymbol{x}_N\,\psi_n^*(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)\psi_m(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N) = \delta_{nm}. \tag{4.6}$$

The non-separable solutions of the TDSE can be expanded as a time-dependent linear combination of the solutions found through separation of variables,

$$\begin{aligned}
\Psi(\boldsymbol{x},t) &= \sum_n c_n(t)\psi_n(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)\phi_n(t), \\
&= \sum_n c_n(t)\psi_n(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)e^{-iE_n(t-t_0)},
\end{aligned} \tag{4.7}$$

where

$$c_n(t_0) = \int \mathrm{d}\boldsymbol{x}_1\cdots\mathrm{d}\boldsymbol{x}_N\,\psi_n^*(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)\Psi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N,t_0). \tag{4.8}$$

4.2   Hartree-Fock method

The Hartree method [27] is a simple method for solving the electronic TISE approximately, which starts by assuming that the N-electron wavefunction $\Psi(\boldsymbol{x})$ can be written as a product of N independent one-electron functions

$$\Psi(\boldsymbol{x}) \approx \psi_1(\boldsymbol{x}_1)\cdots\psi_N(\boldsymbol{x}_N). \tag{4.9}$$

In this approximation, an electron interacts with the other electrons through a mean field, corresponding to the Coulomb repulsion from

the other electrons averaged over their degrees of freedom. To calculate the mean field, an iterative process is often used where an initial guess is refined until a self-consistent solution is obtained. This means that the mean field does not change significantly between iterations. The equation Equation (4.9) used in this approximation excludes the exchange interaction and all correlation between the electrons, including Coulomb correlation. The exchange interaction arises from the Pauli exclusion principle, which forbids identical fermions from occupying the same quantum state and can be expressed mathematically by requiring the wavefunction to be antisymmetric under the exchange of two identical fermions. The electronic Coulomb correlation is the part of the electron-electron Coulomb repulsion that is not accounted for by the mean field.

The Hartree-Fock method [28] improves on the description given by the Hartree method, by assuming that the wavefunction can be written as a Slater determinant

$$\Psi(\boldsymbol{x}) \approx \Psi_{\mathrm{HF}}(\boldsymbol{x})$$

$$= \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\boldsymbol{x}_1) & \psi_2(\boldsymbol{x}_1) & \cdots & \psi_N(\boldsymbol{x}_1) \\ \psi_1(\boldsymbol{x}_2) & \psi_2(\boldsymbol{x}_2) & \cdots & \psi_N(\boldsymbol{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(\boldsymbol{x}_N) & \psi_2(\boldsymbol{x}_N) & \cdots & \psi_{N_e}(\boldsymbol{x}_N), \end{vmatrix} \qquad (4.10)$$

where the N functions $\psi_1, \psi_2, \ldots, \psi_N$ are known as the occupied spin orbitals of the molecule. The Slater determinant formulation gives an exact representation of the electronic exchange correlation, since the determinant is antisymmetric under the exchange of two electrons

$$\Psi_{\mathrm{HF}}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_I, \ldots, \boldsymbol{x}_J, \ldots, \boldsymbol{x}_N) = -\Psi_{\mathrm{HF}}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_J, \ldots, \boldsymbol{x}_I, \ldots, \boldsymbol{x}_N).$$
$$(4.11)$$

Since the position and spin operators commute, each spin orbital can be written as the product $\psi_I(\boldsymbol{x}) = \phi_I(\boldsymbol{r})\chi_I(\boldsymbol{\sigma})$, where $\phi_I(\boldsymbol{r})$ is a spatial orbital and $\chi_I(\boldsymbol{\sigma})$ a spin function. The spin function can be chosen as one of the orthonormal functions $\alpha(\boldsymbol{\sigma})$ and $\beta(\boldsymbol{\sigma})$, corresponding to "spin up" (↑) and "spin down" (↓), respectively, along some quantization axis. Together, $\alpha(\boldsymbol{\sigma})$ and $\beta(\boldsymbol{\sigma})$ form a complete basis for the spin space. The spatial orbitals are expanded in a set of position-dependent mathematical functions $\varphi_J(\boldsymbol{r})$, referred to as the basis set,

$$\phi_I(\boldsymbol{r}) = \sum_J C_{IJ} \varphi_J(\boldsymbol{r}). \qquad (4.12)$$

When the number of functions in the basis set M exceeds the number of electrons N, the additional degrees of freedom lead to $M - N$ virtual

spin orbitals. Although these orbitals do not take part in the self-consistent Hartree-Fock determinant, they play an important role during the Hartree-Fock optimization and for post-Hartree-Fock methods.

In practice, Hartree-Fock calculations are performed using incomplete basis sets. This is a significant source of error, as insufficiently flexible basis sets can fail to properly describe the spatial orbitals obtained in the complete basis set limit. However, the computational cost of Hartree-Fock and post-Hartree-Fock calculations grows quickly with respect to basis set size, requiring a compromise between cost and accuracy. In the simulations of this thesis, the spatial molecular orbitals are expanded in model atomic orbitals centered at the positions of the different nuclei, represented by contracted Gaussian functions.

4.3 Second quantization

The second quantization formalism is often used to simplify the formulation of multiparticle quantum theories, including Hartree-Fock and post Hartree-Fock methods. In this formalism, the occupation of the different spin orbitals can be described by the vector

$$|\boldsymbol{k}\rangle = |k_1, k_2, \ldots k_M\rangle,\qquad(4.13)$$

where M is the number of orbitals, and the occupation number

$$k_P = \begin{cases} 1 & \text{if } \psi_P \text{ is occupied,} \\ 0 & \text{otherwise,} \end{cases}\qquad(4.14)$$

can be modified by the Hermitian conjugate creation $a_P^\dagger$ and annihilation a_P operators. The antisymmetry under fermion exchange can be enforced by letting the operators satisfy the following anticommutation relations

$$\{a_P^\dagger, a_Q^\dagger\} = \{a_P, a_Q\} = 0,\qquad(4.15)$$

$$\{a_P^\dagger, a_Q\} = \delta_{PQ},\qquad(4.16)$$

where $\{A, B\} = AB + BA$. Following Ref. [29], the creation operator $a_P^\dagger$ can be expressed as

$$a_P^\dagger |\boldsymbol{k}\rangle = \delta_{k_P 0} \Gamma_P^{\boldsymbol{k}} |k_1, \ldots, 1_P, \ldots, k_M\rangle,\qquad(4.17)$$

adding an electron to the spin orbital with index P. The Kronecker delta $\delta_{k_P 0}$ ensures that creation of an electron in an occupied spin-orbital is forbidden, and the coefficient $\Gamma_P^{\boldsymbol{k}}$ is the sign given by

$$\Gamma_P^{\boldsymbol{k}} = \prod_{Q=1}^{P-1} (-1)^{k_Q}.\qquad(4.18)$$

The annihilation operator a_P can be expressed as

$$a_P \ket{\boldsymbol{k}} = \delta_{k_P 1} \Gamma_P^k \ket{k_1, \dots, 0_P, \dots, k_M}, \tag{4.19}$$

removing an electron from the spin orbital with index P. The Kronecker delta $\delta_{k_P 1}$ ensures that annihilation of an electron in an unoccupied spin-orbital is forbidden. The Hartree-Fock determinant can be expressed as the vacuum state acted on by the product of the N creation operators corresponding to the lowest-energy orbitals

$$\ket{\text{HF}} = \prod_{I=1}^N a_I^\dagger \ket{0_1, \dots, 0_M}. \tag{4.20}$$

In the position-spin representation, the general one- and two-electron operators f and g can be expressed as

$$f = \sum_i f(\boldsymbol{x}_i), \tag{4.21}$$

$$g = \frac{1}{2} \sum_{i \neq j} g(\boldsymbol{x}_i, \boldsymbol{x}_j). \tag{4.22}$$

where the indices i and j run over the N electrons in the system. In the second-quantization formalism, however, the general one- and two-electron operators f and g can be expressed as

$$f = \sum_{PQ} f_{PQ} a_P^\dagger a_Q, \tag{4.23}$$

$$g = \frac{1}{2} \sum_{PQRS} g_{PQRS} a_P^\dagger a_R^\dagger a_S a_Q, \tag{4.24}$$

respectively, where the one- and two-electron integrals

$$f_{PQ} = \int d\boldsymbol{x} \, \psi_P^*(\boldsymbol{x}) f(\boldsymbol{x}) \psi_Q(\boldsymbol{x}), \tag{4.25}$$

$$g_{PQRS} = \int d\boldsymbol{x} \int d\boldsymbol{x}' \, \psi_P^*(\boldsymbol{x}) \psi_R^*(\boldsymbol{x}') g(\boldsymbol{x}, \boldsymbol{x}') \psi_Q(\boldsymbol{x}) \psi_S(\boldsymbol{x}'), \tag{4.26}$$

respectively.

The field-free Hamiltonian in Equation (2.28) is a two-electron operator, and can in accordance with Equation (4.25) and Equation (4.26) be written as the second-quantization operator

$$H_0 = \sum_{PQ} h_{PQ} a_P^\dagger a_Q + \frac{1}{2} \sum_{PQRS} g_{PQRS} a_P^\dagger a_R^\dagger a_S a_Q + h_{\text{nuc}}, \tag{4.27}$$

where h_{PQ} and $\frac{1}{2}g_{PQRS}$ are the integrals of the one- and two-electron parts of the Hamiltonian, respectively, and the constant nuclear-nuclear Coulomb interaction term is given by

$$h_{\text{nuc}} = \frac{1}{2} \sum_{I \neq J} \frac{Z_I Z_J}{|\boldsymbol{R}_I - \boldsymbol{R}_J|}. \tag{4.28}$$

Note that the two-electron integrals g_{PQRS} are symmetric under particle permutation $PQ \leftrightarrow RS$, and the matrix $\boldsymbol{g}$ with compound indices (PQ, RS) is thus positive definite [30]. This means that $\boldsymbol{g}$ can be decomposed into the product $\boldsymbol{g} = \boldsymbol{LL}^\dagger$ [31] using Cholesky decomposition. This observation can be leveraged to reduce the computational cost of calculations depending on the two-electron integrals [32], which is done throughout the e^T program [33].

Furthermore, we note that the dipole operator taking part in the interaction term of the time-dependent Hamiltonian can be written as the one-electron operator

$$\boldsymbol{d} = \sum_{PQ} \boldsymbol{d}_{PQ} a_P^\dagger a_Q \tag{4.29}$$

in the second quantization formalism.

4.4   Spin restriction

We restrict the Hartree-Fock wavefunction to be an eigenfunction of the squared total spin operator S^2. This is done by restricting the spin orbitals to come in pairs $\psi_{p\alpha}(\boldsymbol{x}) = \phi_p(\boldsymbol{r})\alpha(\boldsymbol{\sigma})$ and $\psi_{p\beta}(\boldsymbol{x}) = \phi_p(\boldsymbol{r})\beta(\boldsymbol{\sigma})$ of a spatial orbital $\phi_p(\boldsymbol{r})$ multiplied with the two spin functions $\alpha(\boldsymbol{\sigma})$ and $\beta(\boldsymbol{\sigma})$. Consequently, the compound spin and spatial index $p\sigma$ is used instead of the spin-orbital index P, giving

$$a_P^\dagger = a_{p\sigma}^\dagger, \tag{4.30}$$

$$a_P = a_{p\sigma}. \tag{4.31}$$

Since the Hamiltonian in Equation (2.28) does not influence the spin of the wavefunction, the spin-orbital indices of the one- and two-electron integrals can be expressed in terms of compound indices, $(P, Q, R, S) = (p\sigma, q\tau, r\mu, s\nu)$, and the integrals rewritten as

$$h_{PQ} = h_{p\sigma q\tau} = h_{pq}\delta_{\sigma\tau}, \tag{4.32}$$

$$g_{PQRS} = g_{p\sigma q\tau r\mu s\nu} = g_{pqrs}\delta_{\sigma\tau}\delta_{\mu\nu}, \tag{4.33}$$

where the spatial integrals

$$h_{pq} = \int d\boldsymbol{r} \, \phi_p^*(\boldsymbol{r}) \left(-\frac{\nabla^2}{2} - \sum_I \frac{Z_I}{|\boldsymbol{r} - \boldsymbol{R}_I|} \right) \phi_q(\boldsymbol{r}), \tag{4.34}$$

$$g_{pqrs} = \int d\boldsymbol{r} \int d\boldsymbol{r}' \, \phi_p^*(\boldsymbol{r}) \phi_r^*(\boldsymbol{r}') \frac{1}{|\boldsymbol{r} - \boldsymbol{r}'|} \phi_q(\boldsymbol{r}) \phi_s(\boldsymbol{r}'). \tag{4.35}$$

Inserting Equations (4.30) to (4.33) into Equation (4.27), we obtain the following expression for the field-free Hamiltonian

$$H_0 = \sum_{pq} h_{pq} E_{pq} + \frac{1}{2} \sum_{pqrs} g_{pqrs} e_{pqrs} + h_{\text{nuc}}, \tag{4.36}$$

where we define the one- and two-electron singlet excitation operators

$$E_{pq} = \sum_\sigma a_{p\sigma}^\dagger a_{q\sigma}, \tag{4.37}$$

$$\begin{aligned} e_{pqrs} &= \sum_{\sigma\tau} a_{p\sigma}^\dagger a_{r\tau}^\dagger a_{s\tau} a_{q\sigma} \\ &= E_{pq} E_{rs} - \delta_{qr} E_{ps}. \end{aligned} \tag{4.38}$$

Furthermore, the dipole moment integrals can in accordance with Equation (4.32) and Equation (4.34) be written as

$$\boldsymbol{d}_{PQ} = \boldsymbol{d}_{pq} \delta_{\sigma\tau}, \tag{4.39}$$

where

$$\boldsymbol{d}_{pq} = -\int d\boldsymbol{r} \, \phi_p^*(\boldsymbol{r}) \boldsymbol{r} \phi_q(\boldsymbol{r}). \tag{4.40}$$

The spin of a singlet wavefunction, which has spin quantum number $s = 0$, is conserved under operation by the singlet excitation operators, as they excite the electrons in a spin orbital pair equally.

We furthermore assume that the Hartree-Fock wavefunction is closed-shell, where the spin orbital pairs are either occupied or empty. This gives the regular restricted Hartree-Fock method, where the total spin of the wavefunction is zero. The left part of Figure 4.1 illustrates how the occupied orbitals, which are the lowest-energy orbitals of the system, are filled with paired electrons in the restricted Hartree-Fock determinant. The center and right parts of the figure illustrates how electrons in the restricted Hartree-Fock determinant can be promoted from the HOMO to the LUMO orbital by one- and two-electron singlet excitation operators.
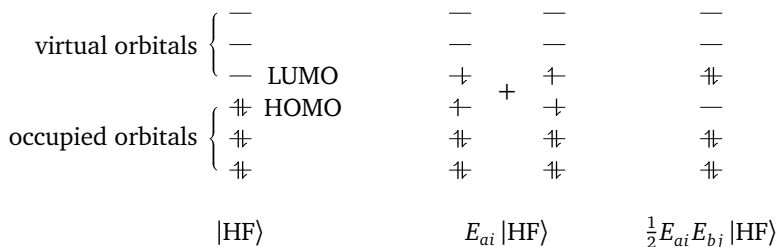
Figure 4.1: Molecular orbital diagrams illustrating how electrons pairs occupy the orbitals of the Hartree-Fock determinant, and how E_{ai} and $\frac{1}{2}E_{ai}E_{bj}$ excite the Hartree-Fock determinant for $a = b = $ LUMO and $i = j = $ HOMO. Note that the excited determinants can be open-shell, but are spin adapted.

4.5   The Roothaan-Hall equations

The average Coulomb potential experienced by each electron from the other electrons in the system is represented by the Fock potential

$$V = \sum_i \left(2g_{pqii} - g_{piiq}\right)E_{pq}. \tag{4.41}$$

The Fock operator f is the sum of the one-electron part of the Hamiltonian and the Fock potential,

$$\begin{aligned} f &= h + V \\ &= \sum_{pq} f_{pq} E_{pq} \end{aligned} \tag{4.42}$$

where

$$f_{pq} = h_{pq} + \sum_i \left(2g_{pqii} - g_{piiq}\right). \tag{4.43}$$

In the basis of the canonical Hartree-Fock molecular orbitals, the Fock operator is diagonal,

$$f_{pq} = \epsilon_p \delta_{pq} \tag{4.44}$$

where ϵ_p is the orbital energy of molecular orbital p. This is however generally not the case, and the canonical Hartree-Fock orbitals can be found by solving the Roothaan-Hall equations [30, 34]

$$fC = SC\epsilon \tag{4.45}$$

iteratively. In these equations, f is the Fock operator expressed in an initial basis and S is the overlap matrix of the initial basis functions. C the matrix
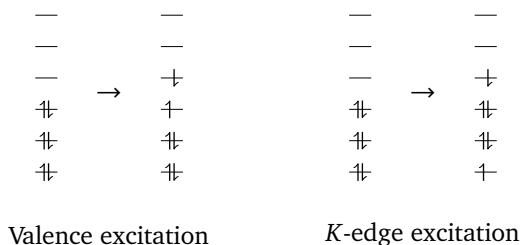
Valence excitation   *K*-edge excitation

Figure 4.2: Molecular orbital diagrams illustrating a valence excitation from the HOMO to the LUMO orbital (left) and a core excitation from the *K*-shell orbital to the LUMO orbital (right).

of orbital coefficients, which relate the initial basis functions to the solution at a given iteration, see Equation (4.12). The equations are nonlinear in the orbital coefficients, which is why an iterative approach is needed.

4.6 Valence and core orbitals

The core orbitals of a Hartree-Fock determinant are the lower-energy orbitals that are mainly localized around the core of the molecule, while the higher energy valence orbitals are more delocalized and participate more readily in bonding. Although more energy is needed for exciting and ionizing core electrons, the probability of resonant excitation of a core electron can be higher in high-energy radiation. This is reflected in molecular absorption cross-sections, which in general increase steeply as the core-exciting energies are approached from below.

The features of the cross-section associated with core excitation are known as absorption edges, and are named with respect to the core orbitals that are excited by the given energy. The *K*-edge is a manifestation of excitations from the *K*-shell, the *L*-edge of the excitations from the *L*-shell, and so on. In Figure 4.2, we illustrate how a HOMO-LUMO valence excitation differs from a *K*-edge excitation from the *K*-shell to the LUMO orbital, disregarding Coulomb correlation. For most elements, the core edges are in the x-ray frequency range, and the correct representation of the edge features are thus important for models of x-ray spectroscopy. The absorption edges of a molecule are closely linked to the absorption edges of its constituent atoms. The absorption edges are also sensitive to electronic density around the nuclei, and thus the chemical environment of the bound atoms. This gives a spectroscopic footprint that can reveal information about the structure of molecules.

4.7 Correlation

As mentioned in Section 4.2, the electron-electron Coulomb repulsion is approximated by a mean-field in the Hartree-Fock method, and the electronic Coulomb correlation thus disregarded. Using the definition of the Fock operator in Equation (4.42), the second-quantization Hamiltonian in Equation (4.36) can be rewritten as

$$H = f + \Phi + h_{\text{nuc}}, \tag{4.46}$$

where the fluctuation potential Φ is

$$\Phi = \frac{1}{2} \sum_{pqrs} g_{pqrs} e_{pqrs} - V. \tag{4.47}$$

The fluctuation potential is disregarded in the Hartree-Fock method, and post-Hartree-Fock methods are methods that aim to represent the fluctuation potential, either exactly or approximately, by correcting the description given by the Hartree-Fock determinant. A conceptually simple way of doing this is through the standard configuration interaction (CI) method, where the state is parametrized as a linear combination of the excited determinants obtained by applying excitation operators to the Hartree-Fock determinant. The spin-adapted full configuration interaction (FCI) method includes singlet excitation operators to all orders,

$$
\begin{aligned}
|\Psi\rangle &= \sum_{\kappa} \tau_{\kappa} c_{\kappa} |\text{HF}\rangle \\
&= |\text{HF}\rangle c_0 + \sum_{ai} E_{ai} |\text{HF}\rangle c_i^a + \sum_{aibj} E_{ai} E_{bj} |\text{HF}\rangle c_{ij}^{ab} + \dots,
\end{aligned} \tag{4.48}
$$

where τ_{κ} denotes a product of singlet excitation operators. The spin-adapted FCI method gives an exact representation of the fluctuation potential for spin-free Hamiltonian operators, and thus also fully represents the electronic correlation in the given basis. The method is exact in the complete basis set limit. However, the FCI method can in practice only be used for calculating the properties of small molecules, as the number of determinants N_{det}, and thus the computational effort needed for the method scales factorially with system size, since the number of determinants is given by

$$N_{\text{det}} = \binom{M}{N} = \frac{M!}{N!(M-N)!} \tag{4.49}$$

where N is the number of electrons and M the number of orbitals. Alternatively, truncated CI methods can be used, in which the expansion in

Equation (4.48) is truncated at a certain order of one-electron excitation operators. These methods reduce the computational effort required to a polynomial scaling with system size

4.8 Size extensivity and intensivity

The proper behavior of observables with the scaling of the size of the quantum system is a central theme of quantum chemistry, and is seen as particularly important for the correct description of large molecular systems. In the FCI limit, properties can be classified into two categories: size-extensive and size-intensive. Size-extensive properties, such as the energy of the system, are additive for subsystems, that is,

$$E_{AB} = E_A + E_B, \tag{4.50}$$

where A and B are two non-interacting systems. Size-intensive properties, on the other hand, such as excitation energies and transition moments [35], do not scale with the number of subsystems. Approximate methods can misrepresent the properties by giving the wrong scaling behavior. Excitation energies of truncated CI methods, for example, are not size-intensive. The favorable size-scaling behavior is one of the most important reasons for why coupled-cluster methods have gained favor as an alternative to configuration interaction methods.

Chapter 5

Coupled-cluster methods

The time-independent and time-dependent methods used in this thesis all
fall within the framework of coupled-cluster methods. In this chapter, the
ground-state coupled-cluster method is introduced, and the various exten-
sions and numerical methods that are used for simulating the interaction
of atoms and molecules with an external field.

5.1 Coupled-cluster ground state

In the traditional coupled-cluster method [36], the ket vector of the ground
state is parametrized as

$$|\psi_0\rangle = e^T |\text{HF}\rangle, \tag{5.1}$$

In the spin-adapted variant, the cluster operator T is a linear combination
of singlet excitation operators of all orders,

$$
\begin{aligned}
T &= \sum_\lambda \tau_\lambda t_\lambda \\
&= t_0 + \sum_{ai} E_{ai} t_i^a + \frac{1}{2} \sum_{ai \geq bj} E_{ai} E_{bj} t_{ij}^{ab} + \dots,
\end{aligned}
\tag{5.2}
$$

where τ_λ denotes a product of singlet excitation operators and the expan-
sion parameters t_λ are known as the cluster amplitudes. Note that the am-
plitude t_0 is often omitted in time-independent coupled cluster methods,
and it does not participate in the ground state equations. The amplitude is,
however, usually included in time-dependent treatments, and we include
it throughout this chapter for a consistent definition of the cluster operator.

As in the CI method, the full coupled cluster method can be approx-
imated by truncating the cluster operator at a given order of the one-
electron excitation operators. The truncation level needed for an accurate

31

description of the system increases with the strength of the correlation of the targeted quantum state. Truncation at the single excitation level gives the coupled-cluster singles (CCS) method, at the double excitation level gives the coupled-cluster doubles (CCSD) method, and so on, with polynomial computational scaling increasing with the truncation level. Using perturbation theory, additional coupled-cluster methods with intermediate computational scaling can be defined, by expressing coupled-cluster equations in orders of the fluctuation potential given by Equation (4.47).

The ground-state solution of the electronic TISE

$$H\,|\psi_0\rangle = E_0\,|\psi_0\rangle \tag{5.3}$$

is in coupled-cluster methods often found by projecting the equation onto $\langle\text{HF}|\,e^{-T}$ and $\langle\widetilde{\mu}|\,e^{-T}$, for $\mu > 0$, where the excited bra determinant

$$\langle\widetilde{\mu}| = \langle\text{HF}|\,\widetilde{\tau}_\mu^\dagger. \tag{5.4}$$

The excited bra determinants are biorthogonal to the Hartree-Fock ket determinant

$$\langle\widetilde{\mu}|\text{HF}\rangle = 0. \tag{5.5}$$

The projection onto $\langle\widetilde{\mu}|\,e^{-T}$ gives the amplitude equations

$$\langle\widetilde{\mu}|\,\bar{H}\,|\text{HF}\rangle = 0, \tag{5.6}$$

where we denote the similarity transformation of an arbitrary operator O by the cluster operator with an overbar,

$$\bar{O} = e^{-T}Oe^T. \tag{5.7}$$

The number of equations in the form of Equation (5.6) is equal to the number of amplitudes in the cluster operator, and the set of equations can be solved for the ground state amplitudes. The equations are often written out explicitly by using the Baker-Campbell-Hausdorff formula

$$e^{-T}Oe^T = O + \frac{1}{2!}[O,T] + \frac{1}{3!}[[O,T],T] + \frac{1}{3!}[[[O,T],T],T] + \ldots, \tag{5.8}$$

which for the Hamiltonian $O = H$ truncates at the fourth term. Since the equations are non-linear in the amplitudes, they are solved with methods suitable for non-linear equations, such as the Newton method.

Assuming that the Hartree-Fock determinant is normalized,

$$\langle\text{HF}|\text{HF}\rangle = 1, \tag{5.9}$$

the projection of the TISE onto $\langle \mathrm{HF}| \, e^{-T}$ gives the ground state energy equation

$$\langle \mathrm{HF}| \bar{H} \, |\mathrm{HF}\rangle = E_0. \tag{5.10}$$

Once the amplitudes have been found from solution of Equation (5.6), the ground state energy follows from Equation (5.10) and Equation (5.8).

In contrast to Equation (5.10), the general expectation value *ansatz*

$$\langle \mathrm{HF}| \bar{O} \, |\mathrm{HF}\rangle \tag{5.11}$$

does not describe a ground state expectation value that satisfies the Hellman-Feynman theorem. However, this limitation can be overcome by treating coupled-cluster theory in the Lagrangian formulation. In this approach, the amplitude equations in Equation (5.6) are introduced as constraints, accompanied by corresponding Lagrange multipliers $\bar{t}_\mu$, resulting in the following Lagrangian

$$L = \langle \mathrm{HF}| \bar{H} \, |\mathrm{HF}\rangle + \sum_\mu \bar{t}_\mu \, \langle \widetilde{\mu}| \bar{H} \, |\mathrm{HF}\rangle. \tag{5.12}$$

The Lagrange multipliers of the ground state can be found by solving the following set of stationarity conditions

$$\langle \mathrm{HF}| \bar{H} \, |\nu\rangle + \sum_\mu \bar{t}_\mu A_{\mu\nu} = 0, \tag{5.13}$$

where the excited ket determinant

$$|\nu\rangle = \tau_\nu \, |\mathrm{HF}\rangle, \tag{5.14}$$

which we take to be biorthogonal to the bra determinants

$$\langle \mathrm{HF}| \nu\rangle = 0, \quad \langle \widetilde{\mu}| \nu\rangle = \delta_{\mu\nu}. \tag{5.15}$$

The coupled-cluster Jacobian

$$A_{\mu\nu} = \langle \widetilde{\mu}| [\bar{H}, \tau_\nu] \, |\mathrm{HF}\rangle, \tag{5.16}$$

is the derivative of the left-hand side of Equation (5.6) with respect to the amplitude t_ν.

The Lagrangian leads to the ground state expectation value expression

$$\langle O \rangle = \langle \widetilde{\psi}_0| O \, |\psi_0\rangle, \tag{5.17}$$

where

$$\langle \widetilde{\psi}_0| = \langle \mathrm{HF}| \, e^{-T} + \sum_\mu \bar{t}_\mu \, \langle \widetilde{\mu}| \, e^{-T}, \tag{5.18}$$

which obeys the Hellman-Feynman theorem. Note that the equation reduces to Equation (5.10) for $O = H$. Ground state expectation values given by Equation (5.17) are equal to full configuration interaction ground state expectation values when the cluster operator is untruncated.

5.2   Equation-of-motion coupled-cluster method

The equation-of-motion coupled-cluster (EOM-CC) method [37] is a way of extending the ground-state coupled-cluster method to excited states. This is done by defining the coupled-cluster ket and bra of EOM-CC state i as

$$|\psi_i\rangle = e^T R_i \,|\text{HF}\rangle, \tag{5.19}$$

$$\langle\widetilde{\psi}_i| = \langle\text{HF}|\,L_i e^{-T}, \tag{5.20}$$

where the linear excitation operators

$$R_i = \sum_\lambda \tau_\lambda r_{\lambda i}, \tag{5.21}$$

$$L_i = \sum_\kappa l_{i\kappa} \widetilde{\tau}_\kappa^\dagger. \tag{5.22}$$

The summation indices κ and λ include the reference determinant with index $\kappa = \lambda = 0$, and are usually truncated at the same excitation level as the cluster operators. In EOM-CC, the cluster operator T is not allowed to relax in calculations of excited states.

In the following, we derive the well-known EOM-CC eigenvalue equations for the parameters of R_i and L_i. Projecting the associated TISEs

$$H R_i e^T \,|\text{HF}\rangle = e^T R_i \,|\text{HF}\rangle\, E_i \tag{5.23}$$

$$\langle\text{HF}|\,L_i e^{-T} H = E_i \langle\text{HF}|\,L_i e^{-T} \tag{5.24}$$

onto $\langle\widetilde{\kappa}|\,e^{-T}$ and $e^T\,|\lambda\rangle$, respectively, we obtain

$$\sum_\lambda H_{\kappa\lambda} r_{\lambda i} = r_{\kappa i} E_i, \tag{5.25}$$

$$\sum_\kappa l_{i\kappa} H_{\kappa\lambda} = E_i l_{i\lambda}, \tag{5.26}$$

where the elements of the Hamiltonian matrix is given by

$$H_{\kappa\lambda} = \langle\widetilde{\kappa}|\bar{H}\,|\lambda\rangle. \tag{5.27}$$

This matrix can be seen to be non-Hermitian due to the similarity transformation of the Hamiltonian operator. Consequently, both the left and right eigenvalue problems are defined, where in general both eigenvectors are needed for the calculation of EOM-CC operator matrix elements, whereas in Hermitian theories only a single eigenvalue problem is required. The Hamiltonian matrix can be written in the following commutator form

$$H_{\kappa\lambda} = \langle\widetilde{\kappa}|\,[\bar{H}, \tau_\lambda]\,|\text{HF}\rangle + \langle\widetilde{\kappa}|\,\tau_\lambda \bar{H}\,|\text{HF}\rangle \tag{5.28}$$

Inserting the resolution of identity $\mathbb{1} = \sum_\eta |\eta\rangle \langle \tilde{\eta}|$ in the last term, and also Equation (5.6) and Equation (5.10), we obtain

$$H_{\kappa\lambda} = \langle \tilde{\kappa}| [\bar{H}, \tau_\lambda] |\text{HF}\rangle + \delta_{\kappa\lambda} E_0 \tag{5.29}$$

Inserting this back into Equation (5.25) and Equation (5.26), we obtain

$$\sum_\nu \langle \tilde{\kappa}| [\bar{H}, \tau_\nu] |\text{HF}\rangle \, r_{\nu i} = r_{\kappa i} \Delta E_i, \tag{5.30}$$

$$\sum_\kappa l_{i\kappa} \langle \tilde{\kappa}| [\bar{H}, \tau_\lambda] |\text{HF}\rangle = \Delta E_i l_{i\lambda}, \tag{5.31}$$

where $\Delta E_i = E_i - E_0$.

For the indices $\kappa = \mu$, where $\mu > 0$, Equation (5.30) reduces to

$$\sum_\nu A_{\mu\nu} r_{\nu i} = r_{\mu i} \Delta E_i, \tag{5.32}$$

which defines the right EOM-CC eigenvalue problem. For $\kappa = 0$, the equation reduces to

$$\sum_\nu \langle \text{HF}| \bar{H} |\nu\rangle \, r_{\nu i} = r_{0i} \Delta E_i, \tag{5.33}$$

which after inserting Equation (5.13) and Equation (5.32) and assuming $\Delta E_i \neq 0$ gives

$$r_{0i} = -\sum_\mu \bar{t}_\mu r_{\mu i}. \tag{5.34}$$

For $\lambda = 0$, Equation (5.31) reduces to

$$\Delta E_i l_{i0} = 0, \tag{5.35}$$

which for $\Delta E_i \neq 0$ gives

$$l_{0i} = 0. \tag{5.36}$$

For $\lambda = \nu$, where $\nu > 0$, the equation reduces to

$$\sum_\kappa l_{i\kappa} \langle \tilde{\kappa}| [\bar{H}, \tau_\nu] |\text{HF}\rangle = \Delta E_i l_{i\nu}, \tag{5.37}$$

which after inserting Equation (5.36) gives

$$\sum_\mu l_{i\mu} A_{\mu\nu} = \Delta E_i l_{i\nu}, \tag{5.38}$$

which defines the left eigenvalue problem.

Equation (5.32) and Equation (5.38) can be written in the matrix form

$$\boldsymbol{AR}_i = \boldsymbol{R}_i \Delta E_i, \tag{5.39}$$

$$\boldsymbol{A}^T \boldsymbol{L}_i = \Delta E_i \boldsymbol{L}_i, \tag{5.40}$$

respectively, where the vectors $\boldsymbol{R}_i$ and $\boldsymbol{L}_i$ contain the excited determinant components $r_{\mu i}$ and $l_{i\nu}$, and the Jacobian matrix $\boldsymbol{A}$ is non-Hermitian. These eigenvalue problems generally have multiple solutions, which come in addition to the single ground-state solution with energy E_0 obtained from the ground state coupled-cluster equations.

In EOM-CC, the operator O can be expressed as a matrix in the basis of the different EOM-CC eigenstates, with elements

$$\begin{aligned} O_{ij} &= \langle \widetilde{\psi}_i | O | \psi_j \rangle \\ &= \langle \mathrm{HF} | L_i \bar{O} R_j | \mathrm{HF} \rangle, \end{aligned} \tag{5.41}$$

where R_j comes from the solution of Equation (5.39) and L_i from the solution of Equation (5.40). For $i = j$, Equation (5.41) reduces to expressions for the EOM-CC expectation values of O.

5.3 Core-valence separation approximation

Core-excited solutions of Equation (5.39) and Equation (5.40) can be difficult to calculate, since these often have high-lying eigenvalues embedded in a valence ionization (pseudo-)continuum. Some eigenvalue algorithms, such as the ones based on the shift-and-invert procedure, can be used to target specific spectral regions, but these generally require the solution of additional linear equations. An alternative approach is to use the core-valence separation (CVS) approximation [37], where the off-diagonal core-valence and valence-core determinant blocks of the Jacobian are neglected. Then, Equation (5.39) and Equation (5.40) separate into independent eigenvalue problems for the valence-valence and core-core determinant blocks, where the lowest-eigenvalue solutions are approximate valence-excited and core-excited states, respectively. In this way, established and well-performing exterior eigenvalue algorithms can also be used for calculating core-excited states.

We calculate the core-excited states in the CVS approximation by projecting out all EOM-CC vector elements that do not reference at least a single core orbital at each iteration of the eigenvalue algorithms, which for EOM-CCSD can be stated as [38]

$$\mathcal{P}_I^{\mathrm{CVS}} r_i^a = l_i^a \mathcal{P}_I^{\mathrm{CVS}} = 0, \quad \text{if } i \notin I, \tag{5.42}$$

$$\mathcal{P}_I^{\mathrm{CVS}} r_{ij}^{ab} = l_{ij}^{ab} \mathcal{P}_I^{\mathrm{CVS}} = 0, \quad \text{if } i \notin I \text{ and } j \notin I. \tag{5.43}$$

This projection effectively removes the valence-valence, core-valence and valence-core determinant blocks of the Jacobian, decoupling the core-core block from the rest of the spectrum. The complementary projector can be used in order to calculate the valence-excited states in the CVS approximation, which by design are biorthogonal to the CVS approximated core states. If the valence-excited states are instead calculated as the lowest-eigenvalue solutions of the full Jacobian, they can have a non-zero overlap with the CVS approximated core-excited states. The Rayleigh-Ritz procedure can then be used in order to generate a non-overlapping set of valence- and core-excited states, which we do in Paper III and Paper V.

5.4 Methods for solving the EOM-CC eigenvalue problem

The full Jacobian matrix A is often too large to be stored in memory or even on disk, as the number of elements is equal to the number of determinants in the projection space squared. These limitations are often circumvented by the use of matrix-free methods, where the matrix is accessed only through the matrix-vector products $\rho = Ar$ and $\sigma = A^T l$, which can be used to construct solutions to the eigenvalue problem. The calculation and storage of all solutions is also often not possible for large systems, and it is thus also important to use methods that target the most important states for describing the physical process of interest.

In the papers in this thesis, we use matrix-free Davidson [39] and Lanczos [40] methods to construct approximate solutions to Equation (5.39) and Equation (5.40). The two methods have many features in common. Both start out from predefined starting vectors, and iteratively generate a basis from these vectors. Both methods also use the Rayleigh-Ritz method in order to approximate the solutions of the eigenvalue problem, by diagonalizing the matrix in the generated basis. The approximate solutions generally converge towards the exact solutions as the size of the basis increases, but this convergence is also sensitive to the choice of the starting vectors, which should be chosen with respect to the solutions of interest. The iteratively generated basis vectors often start to overlap as the algorithms proceed, and therefore orthogonalization procedures are often used to ensure the stability of the methods. Both methods perform well at converging the outer eigenvalues of a spectrum, provided they are well-separated.

The Lanczos and Davidson methods do, however, have some distinctive characteristics. In the asymmetric Lanczos method [41], the expan-

sion of the basis is done through repeated multiplication of the right and left starting vectors by the matrices A and A^T, respectively, meaning that the method is a Krylov subspace method. More details about the Lanczos algorithm that we have used are given in Paper III. In each iteration of the Davidson method, the Rayleigh-Ritz method is used to construct the approximate right and left solutions. Then, the basis is expanded by applying the inverse preconditioner

$$P^{-1} = (D - \widetilde{\omega}_i I)^{-1}, \tag{5.44}$$

to the residuals of each insufficiently converged solution, where $\widetilde{\omega}_i$ is its approximate eigenvalue and D is a diagonal approximation of the matrix A. In Paper II, we use the diagonal approximation

$$D_{\mu\nu} = \epsilon_\mu \delta_{\mu\nu}, \tag{5.45}$$

which is equal to the term originating from the Fock operator in the Jacobian, where ϵ_μ is the orbital energy differences of excited determinant μ with respect to the reference Hartree-Fock determinant.

When compared to Krylov space methods, the construction of the approximate solutions and preconditioning at each iteration adds significant cost per iteration for the Davidson method. However, for diagonally dominant matrices, which are commonly encountered in quantum chemistry, the Davidson method has an asymptotic quadratic convergence towards specific eigenvalues. This is since the preconditioning can increase the separation of the targeted eigenvalues from the rest of the spectrum [42]. Note, however, that the method breaks down for eigenvalue problems with completely diagonal matrices, as the generated basis vectors are linearly dependent on the previous ones. The Lanczos method tends to have a slower convergence towards specific eigenvalues, but on other hand often gives a better global convergence, and does not break down for diagonal matrices.

5.5 Time-dependent coupled-cluster method

The time-dependent coupled cluster method [43] can be derived by projecting the time-dependent Schrödinger equations for the time-dependent ket $|\Psi(t)\rangle = e^{T(t)} |\text{HF}\rangle$ and bra $\langle \widetilde{\Psi}(t)| = \sum_\kappa \bar{t}_\kappa(t) \langle \widetilde{\kappa}| e^{-T(t)}$, namely

$$i\frac{\mathrm{d}}{\mathrm{d}t}\big(e^{T(t)}|\text{HF}\rangle\big) = H(t)e^{T(t)}|\text{HF}\rangle, \tag{5.46}$$

$$-i\frac{\mathrm{d}}{\mathrm{d}t}\bigg(\sum_\kappa \bar{t}_\kappa(t)\langle\widetilde{\kappa}|e^{-T(t)}\bigg) = \sum_\kappa \bar{t}_\kappa(t)\langle\widetilde{\kappa}|e^{-T(t)}H(t), \tag{5.47}$$

onto $\langle \widetilde{\kappa} | e^{-T(t)}$ and $e^{T(t)} | \lambda \rangle$, respectively, giving

$$i\frac{\mathrm{d}t_\kappa(t)}{\mathrm{d}t} = \langle \widetilde{\kappa} | \bar{H}(t) | \mathrm{HF} \rangle , \tag{5.48}$$

$$-i\frac{\mathrm{d}\bar{t}_\lambda(t)}{\mathrm{d}t} = \sum_\kappa \bar{t}_\kappa \langle \widetilde{\kappa} | [\bar{H}(t), \tau_\lambda] | \mathrm{HF} \rangle . \tag{5.49}$$

The structure of the TDCC equations ensure that expectation values given by Equation (5.17) scale properly with the size of the system, see Paper IV.

The coupled-cluster response method is a frequency-space method based on Equation (5.48) and Equation (5.49). The method can be derived by expressing the parameters in terms of the orders of perturbation,

$$t_\kappa(t) = t_\kappa^{(0)}(t) + t_\kappa^{(1)}(t) + t_\kappa^{(2)}(t) + \dots \tag{5.50}$$

$$\bar{t}_\lambda(t) = \bar{t}_\lambda^{(0)}(t) + \bar{t}_\lambda^{(1)}(t) + \bar{t}_\lambda^{(2)}(t) + \dots, \tag{5.51}$$

and inserting the Fourier transform of the time-dependent fields, resulting in expressions for frequency-dependent response functions [43]. The method gives size-intensive molecular properties. It does, however, require the solution of extra sets of linear equations, and implementations are non-trivial for higher-order perturbations.

In addition to the perturbative approach given by the coupled-cluster response method, the set of nonlinear equations given by Equation (5.48) and Equation (5.49) can be solved in the real time domain [14], which is the approach used in this thesis.

5.6 Time-dependent equation-of-motion coupled-cluster method

In the time-dependent equation-of-motion coupled-cluster (TD-EOM-CC) method [15, 16], the time-dependence of the state is parametrized by the linear singlet excitation operator $R(t)$ and deexcitation operator $L(t)$, while the cluster operator T is taken to be constant. As in TDCC, the time dependence of the TD-EOM-CC parameters can be found by projecting the time-dependent Schrödinger equations of the ket $|\Psi(t)\rangle = e^T R(t) |\mathrm{HF}\rangle$ and bra $\langle \widetilde{\Psi}(t)| = \langle \mathrm{HF}| L(t)e^{-T}$

$$i\frac{\mathrm{d}}{\mathrm{d}t}\Big(e^T R(t) |\mathrm{HF}\rangle \Big) = H(t)e^T R(t) |\mathrm{HF}\rangle , \tag{5.52}$$

$$-i\frac{\mathrm{d}}{\mathrm{d}t}\Big(\langle \mathrm{HF}| L(t)e^{-T} \Big) = \langle \mathrm{HF}| L(t)e^{-T}H(t), \tag{5.53}$$

onto $\langle \widetilde{\kappa} | e^{-T(t)}$ and $e^{T(t)} | \lambda \rangle$, respectively, giving the equations

$$i \frac{\mathrm{d}r_\kappa(t)}{\mathrm{d}t} = \sum_\lambda \langle \widetilde{\kappa} | \bar{H}(t) | \lambda \rangle \, r_\lambda(t), \qquad (5.54)$$

$$-i \frac{\mathrm{d}l_\lambda(t)}{\mathrm{d}t} = \sum_\kappa l_\kappa(t) \langle \widetilde{\kappa} | \bar{H}(t) | \lambda \rangle. \qquad (5.55)$$

Time-dependent expectation values can be found through the expression

$$\begin{aligned} \langle O(t) \rangle &= \langle \widetilde{\Psi}(t) | O(t) | \Psi(t) \rangle \\ &= \sum_{\kappa\lambda} l_\kappa(t) \langle \widetilde{\kappa} | \bar{O}(t) | \lambda \rangle \, r_\lambda(t). \end{aligned} \qquad (5.56)$$

The TD-EOM-CC method shares similarities with the time-dependent configuration-interaction (TD-CI) method, where the time dependence is also parametrized by linear operators. The CI method does, however, not incorporate the similarity transformation used in coupled cluster methods, and as a result methods based on the truncated CI parametrization do not exhibit the correct scaling of expectation values and transition moments with the size of the system. Conversely, ground state expectation values and excitation energies are correctly represented in truncated EOM-CC methods. In contrast to coupled cluster response theory, the EOM-CC frequency response can be deduced by inserting the biorthogonal EOM-CC states into the expressions obtained for exact states [44]. The equations describing the time evolution of the TD-EOM-CC state are also linear in the time-dependent parameters, while the equations for the time-dependent TDCC amplitudes are non-linear.

In this thesis, Equation (5.54) and Equation (5.55) are solved in the real time domain, both directly in the elementary basis given by Equation (5.54) and Equation (5.55) and also in the eigenstate basis obtained by diagonalizing the field-free Hamiltonian, in which the time-dependent state can be expressed as

$$| \Psi(t) \rangle = \sum_j | \psi_j \rangle \, c_j(t) \qquad (5.57)$$

$$\langle \widetilde{\Psi}(t) | = \sum_i b_i(t) \langle \widetilde{\psi}_i |. \qquad (5.58)$$

The expectation value expression in Equation (5.56) can also be written in this basis

$$\langle O \rangle = \sum_{ij} b_i(t) O_{ij}(t) c_j(t), \qquad (5.59)$$

where the operator matrices $O_{ij}(t)$ are given by Equation (5.41).

The TD-EOM-CC formulation employed in this thesis is closely related to the approaches presented by Sonk et al. [15] and Luppi and Head-Gordon [16], both of which utilize the field-free eigenstate basis. However, our method differs from these approaches by propagating the state using the full non-Hermitian Hamiltonian without symmetrization, which better aligns it with prevalent non-Hermitian EOM-CC and TDCC formulations. This comes at the cost of allowing the time-dependent energy to become complex.

There are alternative moment-based formulations of TD-EOM-CC as well, in which moment functions are constructed by operating on the bra and ket of the coupled-cluster ground state with a one-electron operator of interest, such as the dipole moment operator [45–48]. In these approaches, one of the moment functions is propagated using the time-independent electronic Hamiltonian. To simulate linear electron spectroscopy, the auto-correlation function is employed, involving the product of the propagated and unpropagated moment functions. However, this method does not account for nonlinearities, making it unsuitable for simulating multiphoton processes, which are the primary focus of this thesis.

In Paper III, we demonstrate that TD-EOM-CC can give reasonable results for small molecular systems in weak fields. Properties other than the energy can, however, have the wrong size-scaling behavior. In Paper IV, we demonstrate that the errors of TD-EOM-CC get more pronounced in strong fields, in particular for systems containing multiple subsystems that are in resonance with an external field.

5.7   Real-time integration algorithms

In the TDCC and TD-EOM-CC methods, the system of equations obtained from projecting the right and left TDSEs can be written as

$$\frac{\mathrm{d}\boldsymbol{y}(t)}{\mathrm{d}t} = \boldsymbol{f}(t, \boldsymbol{y}(t)), \qquad (5.60)$$

where the vector $\boldsymbol{y}(t)$ contains all the time-dependent parameters of the method. The TD-EOM-CC equations are linear in $\boldsymbol{y}(t)$, while the TDCC ones are nonlinear. Assuming that the parameters are known at $t = t_0$, the problem of finding the parameters at a later time is an initial value problem, which can be solved by integrating Equation (5.60) numerically in discrete time steps. Note that in addition to the implicit time dependence of the parameters, Equation (5.60) also allows for an explicit dependence on time, which enters for interactions with a time-dependent external field.

In all papers in this thesis, Equation (5.60) is integrated with Runge-Kutta methods. These methods propagate the solution $\boldsymbol{y}(t)$ sequentially in single time steps, and can from step n to $n+1$ be written in the general form [49]

$$\boldsymbol{y}_{n+1} = \boldsymbol{y}_n + h \sum_{i=1}^{\nu} b_i \boldsymbol{f}\left(t_n + c_i h, \boldsymbol{\xi}_i\right) \tag{5.61}$$

$$\boldsymbol{\xi}_i = \boldsymbol{y}_n + h \sum_{j=1}^{\nu} a_{ij} \boldsymbol{f}\left(t_n + c_j h, \boldsymbol{\xi}_j\right), \tag{5.62}$$

where h is the time step size, and the specific method is determined by the matrix $\boldsymbol{a}$ and the vectors $\boldsymbol{b}$ and $\boldsymbol{c}$. The integer ν gives the number of stages per time step. For well-designed Runge-Kutta methods, the number of stages generally correlates with the error of the method, which is classified with respect to the order of the time step size, $\mathcal{O}(h^n)$. The methods are also classified as explicit when all stages are given explicitly in terms of previously calculated stages, and implicit when the stages have an implicit dependence on stages that have not been calculated yet. We explored the use of both explicit and implicit Runge-Kutta methods with fixed time steps in Paper II, and obtained solutions that were well-converged with respect to the time step size for all methods. In Paper III and Paper IV, we used embedded explicit Runge-Kutta methods with adaptive time steps. The time steps were adapted by calculating the size of an error estimate, obtained by evaluating the norm $\|\boldsymbol{y}_n - \boldsymbol{y}_n'\|$ of the difference between the two vectors $\boldsymbol{y}_n$ and $\boldsymbol{y}_n'$ calculated with Runge-Kutta methods of different error orders. These methods are simpler to use, as inadequate choices of the time step size can be automatically corrected for, such as when the integration reaches a region with a large estimated integration error. In Paper III, we show that results calculated with the embedded Runge-Kutta methods indeed reproduce the results of Paper II for large initial time step sizes.

Chapter 6

Select topics in semiclassical laser-molecule interaction

In this chapter, we introduce some concepts and models that help explain the observations made in Papers II to V. We start by introducing the interaction picture and the Dyson series, and use these concepts when introducing the phenomena of symmetry in multiphoton transitions, quantum beats in transient absorption and semiclassical Rabi oscillations.

6.1 Interaction picture and Dyson series

The interaction picture is useful for describing interactions of atoms and molecules with external fields, as the interaction picture Hamiltonian is zero when the interaction is zero. The time-dependent interaction picture ket can be related to the ket in the Schrödinger picture by the unitary transformation

$$|\Psi^{I}(t)\rangle = e^{iH_0 t} |\Psi(t)\rangle, \qquad (6.1)$$

where H_0 is the field-free term of the Hamiltonian. The interaction picture transformation leads to the modified TDSE

$$\frac{\mathrm{d}}{\mathrm{d}t} |\Psi^{I}(t)\rangle = V^{I}(t) |\Psi^{I}(t)\rangle \qquad (6.2)$$

where $V^{I}(t) = e^{iH_0 t} V(t) e^{-iH_0 t}$ and $V(t)$ is the interaction term of the Hamiltonian.

In the interaction picture, the time-evolution operator is related to the Schrödinger picture operator through the transformation

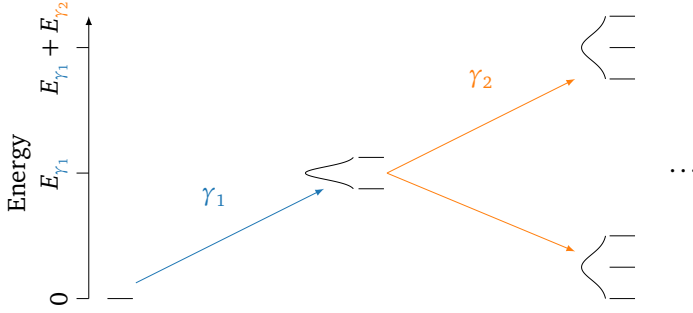$$U^{I}(t, t') = e^{iH_0 t} U(t, t') e^{-iH_0 t'}, \qquad (6.3)$$

Figure 6.1: Multiphoton transitions. From the ground state, the system can absorb a single photon from the external field, exciting the system to a higher energy state. Thereafter, the system can be excited or deexcited by stimulated absorption or emission of a second photon. In very strong fields, this process can continue and involve a high number of photons.

and can be written as a Dyson series [50] in increasing orders of $V^{\mathrm{I}}(t)$,

$$U^{\mathrm{I}}(t, t') = \sum_{n \geq 0} U_n^{\mathrm{I}}(t, t'), \tag{6.4}$$

where

$$U_0^{\mathrm{I}}(t, t') = 1 \tag{6.5}$$

$$U_1^{\mathrm{I}}(t, t') = -i \int_{t'}^{t} \mathrm{d}t_1 \, V^{\mathrm{I}}(t_1) \tag{6.6}$$

$$\vdots \tag{6.7}$$

$$U_n^{\mathrm{I}}(t, t') = (-i)^n \int_{t'}^{t} \mathrm{d}t_1 \int_{t'}^{t_1} \mathrm{d}t_2 \cdots \int_{t'}^{t_{n-1}} \mathrm{d}t_n \, V^{\mathrm{I}}(t_1) V^{\mathrm{I}}(t_2) \cdots V^{\mathrm{I}}(t_n). \tag{6.8}$$

The nth order term in the Dyson series can be interpreted as corresponding to a n-photon transition, where each single-photon exchange is associated with the transition between two states of the system. See Figure 6.1 for an illustration of such multiphoton transitions starting from the ground state. The number of significant terms in the expansion depends on the strength of the interaction, and only the first orders are significant for perturbative interactions.

In the basis of the field-free eigenstates, the matrix elements of the interaction term can in the interaction picture, be written as

$$
\begin{aligned}
V_{ij}^{\mathrm{I}}(t) &= \langle \psi_i | e^{iH_0 t} V(t) e^{-iH_0 t} | \psi_j \rangle \\
&= V_{ij}(t) e^{i\Delta E_{ij} t}
\end{aligned} \tag{6.9}
$$

where $\Delta E_{ij} = E_i - E_j$ is the difference between the energies of the field-free energies. The time evolution of the coefficients c_j is thus given by

$$
\begin{aligned}
i\frac{\mathrm{d}c_i(t)}{\mathrm{d}t} &= \sum_j V_{ij}^I(t)c_j(t) \\
&= \sum_j V_{ij}(t)e^{i\Delta E_{ij}t}c_j(t)
\end{aligned}
\tag{6.10}
$$

6.2   Symmetry in multiphoton transitions

The symmetry of atoms or molecules interacting with an external field can provide valuable insights into the possible electronic transitions that may occur within them. This symmetry is commonly explained within the context of group theory, where the atom or molecule and the field are represented by a point group that encompasses all the symmetry operations that the entity is invariant to. Atoms and small molecules are often highly symmetric, and their point groups can therefore contain a high number of symmetry operations, while the number generally decreases with increasing molecular complexity. The diatomic lithium fluoride molecule is for instance invariant to the identity operation, any rotation around the vertical axis, and an infinite number of vertical mirror planes, and is identified with the $C_{\infty v}$ point group. The lowest-energy conformer of the amino acid glycine, however, is only invariant to the reflection through a single mirror plane in addition to the identity operation, and is identified with the point group C_s.

   The representation of a point group in a given basis can be decomposed into a direct sum of irreducible representations, which are representations that cannot be decomposed further and behave differently with respect to the set of symmetry operations of the point group. Electronic eigenstates are often classified in terms of the irreducible representations of the atom or molecule, as each eigenstate can be shown to transform as one of these representations. Qualitative information about the interaction matrix element $V_{ij}(t) = \langle \Psi_i | V(t) | \Psi_j \rangle$ can be deduced from the direct product of three irreducible representations

$$
\Gamma_i \times \Gamma_V \times \Gamma_j,
\tag{6.11}
$$

where eigenstates i and j transform as Γ_i and Γ_j, respectively, and the length gauge interaction operator $V(t) = -\boldsymbol{d} \cdot \boldsymbol{\mathcal{E}}(t)$ is assumed to transform as Γ_V. If the product does not contain the totally symmetric irreducible representation of the point group, Γ_{sym}, the interaction matrix element is zero,

and the electronic transition forbidden. If the product contains Γ_{sym}, on the other hand, the electronic transition is allowed. This does, however, not imply that the transition is important in simulations, as its strength can be negligible and its energy outside the bandwidth of the external field.

So far, we have only examined the symmetry properties of the interaction matrix elements. However, the complete time-dependent interaction is represented by the time evolution operator, which can involve higher-order products of these matrix elements, see Equation (6.4). By using the same approach as we did when describing the interaction matrix elements in terms of irreducible representations, we can infer that an N-photon transition is only allowed when Γ_{sym} is contained in the direct product

$$\Gamma_i \times \underbrace{\Gamma_V \times \cdots \times \Gamma_V}_{N} \times \Gamma_j. \tag{6.12}$$

6.3   Quantum beats in transient absorption

We now outline an explanatory model of quantum beats in transient absorption, which we observe in the simulated transient absorption spectra in Paper II and Paper III. We start by assuming that the interaction term $V(t)$ in Equation (3.29) is a sum of interaction terms corresponding to temporally separated pump $V_1(t)$ and probe $V_2(t)$ pulses

$$V(t) = V_1(t) + V_2(t). \tag{6.13}$$

Both pulses are treated in the dipole approximation and length gauge. We also assume that the pulses are linearly polarized, meaning that their electric field vectors can be written as products of a polarization vector $\hat{\boldsymbol{e}}$ and a time-dependent amplitude $\mathcal{E}(t)$. The terms for the two pulses can thus be written as

$$V_1(t) = -\boldsymbol{d} \cdot \hat{\boldsymbol{\epsilon}}_1 \mathcal{E}_1(t), \tag{6.14}$$
$$V_2(t) = -\boldsymbol{d} \cdot \hat{\boldsymbol{\epsilon}}_2 \mathcal{E}_2(t). \tag{6.15}$$

The amplitude of transition between the initial state i at time t_0 and the final state f at time t, where both states are eigenstates of H_0 with eigenvalues ω_i and ω_f, respectively, is given by

$$
\begin{aligned}
c_{fi}(t) &= \langle \psi_f | U(t, t_0) | \psi_i \rangle \\
&= \langle \psi_f | e^{-iH_0 t} U^{\text{I}}(t, t_0) e^{iH_0 t_0} | \psi_i \rangle \\
&= e^{i(\omega_i t_0 - \omega_f t)} \langle \psi_f | U^{\text{I}}(t, t_0) | \psi_i \rangle.
\end{aligned} \tag{6.16}
$$

Here, $U^{\mathrm{I}}(t, t_0)$ is the interaction picture time evolution operator in Equation (6.3). For simplicity, we assume that that all transitions to state f must involve the absorption of two photons, the first from the pump and the second from the probe. From the second-order term in the Dyson series in Equation (6.4), we then obtain

$$c_{fi}(t) = -e^{i(\omega_i t_0 - \omega_f t)} \int_{t_0}^{t} \mathrm{d}t_1 \int_{t_0}^{t_1} \mathrm{d}t_2 \, \langle \psi_f | V_2^{\mathrm{I}}(t_1) V_1^{\mathrm{I}}(t_2) | \psi_i \rangle \qquad (6.17)$$

We insert the resolution of identity in terms of all states $1 = \sum_n |\psi_n\rangle \langle \psi_n|$, write the interaction picture operators in terms of the Schrödinger picture operators, $V^{\mathrm{I}}(t) = e^{iH_0 t} V(t) e^{-iH_0 t}$, and insert Equations (6.14) to (6.15), obtaining

$$c_{fi}(t) = -e^{i(\omega_i t_0 - \omega_f t)} \sum_n (\boldsymbol{d}_{fn} \cdot \hat{\boldsymbol{\epsilon}}_2)(\boldsymbol{d}_{ni} \cdot \hat{\boldsymbol{\epsilon}}_1)$$

$$\times \int_{t_0}^{t} \mathrm{d}t_1 \, \mathcal{E}_2(t_1) e^{i\omega_{fn} t_1} \int_{t_0}^{t_1} \mathrm{d}t_2 \, \mathcal{E}_1(t_2) e^{i\omega_{ni} t_2},$$

$$(6.18)$$

where $\omega_{ij} = \omega_i - \omega_j$ and the dipole matrix element $\boldsymbol{d}_{ij} = \langle \psi_i | \boldsymbol{d} | \psi_j \rangle$.

We assume that both pulses are temporally confined, implying that they have a non-zero bandwidth in the frequency domain. We further assume that all interactions with the pulses start after time t_0 and end before t, and that the pump interaction ends before the start of the probe interaction. The two integrals then become independent, and their lower and upper limits can be expanded to negative and positive infinity, respectively. We also separate out the effect of the pump-probe time delay τ by describing the probe pulse in terms of the temporally shifted electric field $\mathcal{E}_2^{\tau}(t) = \mathcal{E}_2(t + \tau)$ [51], giving

$$c_{fi}(t) = -e^{i(\omega_i t_0 - \omega_f t)} \sum_n (\boldsymbol{d}_{fn} \cdot \hat{\boldsymbol{\epsilon}}_2)(\boldsymbol{d}_{ni} \cdot \hat{\boldsymbol{\epsilon}}_1)$$

$$\times \int_{-\infty}^{\infty} \mathrm{d}t_1 \, \mathcal{E}_2^{\tau}(t_1 - \tau) e^{i\omega_{fn} t_1} \int_{-\infty}^{\infty} \mathrm{d}t_2 \, \mathcal{E}_1(t_2) e^{i\omega_{ni} t_2}$$

$$(6.19)$$

By substituting $t_1 = t_1' + \tau$, we obtain

$$c_{fi}(t) = -e^{i\omega_i t_0 - i\omega_f(t - \tau)} \sum_n (\boldsymbol{d}_{fn} \cdot \hat{\boldsymbol{\epsilon}}_2)(\boldsymbol{d}_{ni} \cdot \hat{\boldsymbol{\epsilon}}_1) e^{-i\omega_n \tau}$$

$$\times \int_{-\infty}^{\infty} \mathrm{d}t_1 \, \mathcal{E}_2^{\tau}(t_1') e^{i\omega_{fn} t_1'} \int_{-\infty}^{\infty} \mathrm{d}t_2 \, \mathcal{E}_1(t_2) e^{i\omega_{ni} t_2}.$$

$$(6.20)$$

In comparison with the following convention for the Fourier transform,

$$\mathcal{E}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\omega\, \widetilde{\mathcal{E}}(\omega) e^{-i\omega t}, \tag{6.21}$$

$$\widetilde{\mathcal{E}}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dt\, \mathcal{E}(t) e^{i\omega t}, \tag{6.22}$$

we recognize that Equation (6.20) involves the product of two Fourier transform integrals, and can be written as

$$c_{fi}(t) = -2\pi e^{i\omega_i t_0 - i\omega_f(t-\tau)} \sum_n (\boldsymbol{d}_{fn} \cdot \hat{\boldsymbol{\epsilon}}_2)(\boldsymbol{d}_{ni} \cdot \hat{\boldsymbol{\epsilon}}_1) e^{-i\omega_n \tau} \widetilde{\mathcal{E}}_2^{\tau}(\omega_{fn}) \widetilde{\mathcal{E}}_1(\omega_{ni}) \tag{6.23}$$

By defining the real numbers

$$a_{ni} = \sqrt{2\pi}(\boldsymbol{d}_{ni} \cdot \hat{\boldsymbol{\epsilon}}_1) \big|\widetilde{\mathcal{E}}_1(\omega_{ni})\big|, \tag{6.24}$$

$$b_{fn} = \sqrt{2\pi}(\boldsymbol{d}_{fn} \cdot \hat{\boldsymbol{\epsilon}}_2) \big|\widetilde{\mathcal{E}}_2^{\tau}(\omega_{fn})\big|, \tag{6.25}$$

Equation (6.23) can be written as

$$c_{fi}(t) = -e^{i\omega_i t_0 - i\omega_f(t-\tau)} \sum_n b_{fn} a_{ni} e^{-i(\omega_n \tau + \varphi_2(\omega_{fn}) + \varphi_1(\omega_{ni}))} \tag{6.26}$$

where the spectral phases $\varphi_1(\omega) = \arg\big(\widetilde{\mathcal{E}}_1(\omega)\big)$ and $\varphi_2(\omega) = \arg\big(\widetilde{\mathcal{E}}_2^{\tau}(\omega)\big)$.

The probability of transition from state i to state f is equal to the squared magnitude of the transition amplitude

$$\begin{aligned} |c_{fi}|^2 &= \left(\sum_n b_{fn} a_{ni} e^{i(\omega_n \tau + \varphi_2(\omega_{fn}) + \varphi_1(\omega_{ni}))} \right) \\ &\quad \times \left(\sum_m b_{fm} a_{mi} e^{-i(\omega_m \tau + \varphi_2(\omega_{fm}) + \varphi_1(\omega_{mi}))} \right) \\ &= \sum_n b_{fn}^2 a_{ni}^2 + 2 \sum_{n>m} b_{fn} a_{ni} b_{fm} a_{mi} \cos\big(\omega_{nm}\tau + \varphi_{fnmi}\big) \end{aligned} \tag{6.27}$$

where $\varphi_{fnmi} = \varphi_2(\omega_{fn}) - \varphi_2(\omega_{fm}) + \varphi_1(\omega_{ni}) - \varphi_1(\omega_{mi})$. Equivalent expressions have been reported in several earlier works [52–55]. Note that the relative contribution from the different states depends both on the characteristics of the laser pulse and on the transition matrix elements.

In a four state system, the probability of transition from the ground state $i = 0$ to the final state $f = 3$ through the two intermediate states $n = 1, 2$ is given by

$$|c_3|^2 = b_{31}^2 a_{10}^2 + b_{32}^2 a_{20}^2 + 2 b_{32} a_{20} b_{31} a_{10} \cos(\omega_{21}\tau + \varphi_{3210}) \tag{6.28}$$
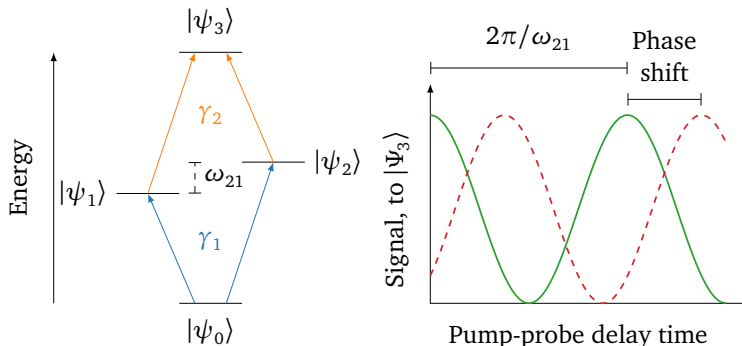
Figure 6.2: Illustration of the quantum beats described by Equation (6.28), occurring in a system with the four states $|\psi_0\rangle$, $|\psi_1\rangle$, $|\psi_2\rangle$ and $|\psi_3\rangle$ and sequential absorption of the photons γ_1 and γ_2. In the left part of the illustration, it is shown how the system can be excited from the initial state $|\psi_0\rangle$ to the final state $|\psi_3\rangle$ through two interfering pathways, where the intermediate states $|\psi_1\rangle$ and $|\psi_2\rangle$ have the energy difference ω_{21}. In the right part of the figure, the absorption signal corresponding to the interfering transitions to $|\psi_3\rangle$ is illustrated, which has the period $2\pi/\omega_{21}$ and a phase shift arising from the spectral phase difference $\varphi_{3210} = \varphi_2(\omega_{32}) - \varphi_2(\omega_{31}) + \varphi_1(\omega_{20}) - \varphi_1(\omega_{10})$.

An illustration of the behavior of the four-state system described by this equation is given in Figure 6.2. Note that the signal changes as a phase shifted cosine function, with the frequency ω_{21} corresponding to the energy difference of the two intermediate states. That is, the signal will display a quantum beat as the two transition pathways interfere in a time-dependent fashion. This is analogous to the double-slit experiment, where the pathways through the two different slits interfere, creating a position-dependent interference pattern.

The overlap between the pump and probe pulses was neglected in the above model of quantum beats. For short pulse delays, this overlap can modify the absorption signal, generating what is known as coherent artifacts. We observe these artifacts in simulated transient absorption spectra in both Paper II and Paper III, where the overlap is not neglected.

Note that Equation (6.27) assumes a precise knowledge about the spectral phases of the pump and probe pulses and the inter-pulse delay. This level of control is unrealistic for many experimental setups, as state-of-the-art free-electron lasers, amplified by self-amplified spontaneous emission (SASE) have stochastic phases. In addition, timing jitter often makes it hard to synchronize the pump and probe pulses [56], and it is difficult to

lower this jitter to values shorter than the typical timescales of electronic motion. However, approaches such as ghost imaging may be used to characterize the pulses in a non-invasive way [57], and for high repetition rate sources, the results can be sorted by pulse characteristics for comparison with theoretical results [58].

6.4   Semiclassical Rabi model

The semiclassical Rabi model gives an explanation of how a two-level quantum system can sequentially undergo stimulated absorption and emission in a classical external field. In the following, we show a simple derivation of the model, following Ref. [59].

We start by assuming that the external laser field is monochromatic, having the frequency ω and the peak electric field in the polarization direction $\mathcal{E}_0$, and write the interaction term in Equation (3.29) as

$$V(t) = \boldsymbol{d} \cdot \mathcal{E}_0 \cos(\omega t). \tag{6.29}$$

We also assume that the system can be described by two eigenstates of the field-free Hamiltonian

$$H_0 \ket{\psi_1} = \omega_1 \ket{\psi_1}, \tag{6.30}$$

$$H_0 \ket{\psi_2} = \omega_2 \ket{\psi_2}, \tag{6.31}$$

and that the diagonal elements of the interaction Hamiltonian are zero,

$$V_{11}(t) = V_{22}(t) = 0, \tag{6.32}$$

which is valid for states with definite parity, such as for all atomic eigenstates. Under these assumptions, the time evolution of the eigenstate coefficients, given by Equation (6.10), reduces to the two equations

$$i \frac{\mathrm{d}c_1(t)}{\mathrm{d}t} = V_{12}(t) e^{i\omega_{12}t} c_2(t), \tag{6.33}$$

$$i \frac{\mathrm{d}c_2(t)}{\mathrm{d}t} = V_{21}(t) e^{i\omega_{21}t} c_1(t). \tag{6.34}$$

We now define the Rabi frequency

$$\Omega = \boldsymbol{d}_{12} \cdot \mathcal{E}_0 = \boldsymbol{d}_{21} \cdot \mathcal{E}_0, \tag{6.35}$$

where the real observable $\boldsymbol{d}_{12} = \boldsymbol{d}_{21}^* = \boldsymbol{d}_{21}$. This gives

$$\begin{aligned} V_{12}(t) = V_{21}(t) \\ = \Omega \cos(\omega t) \\ = \frac{1}{2} \Omega \left(e^{i\omega t} + e^{-i\omega t} \right). \end{aligned} \tag{6.36}$$
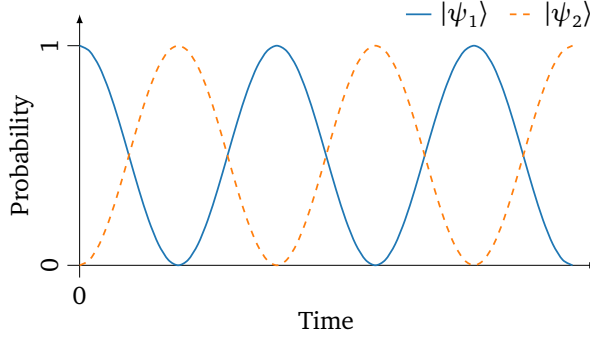
Figure 6.3: Illustration of Rabi oscillations of a two-state system in a monochromatic electric field, where the field is tuned in resonance with the interstate transition. The time-dependent probability of observing the system in state $|\psi_1\rangle$ and $|\psi_2\rangle$ are shown in solid blue and dashed orange lines, respectively.

Inserting Equation (6.36) into Equation (6.33) and Equation (6.34), we obtain

$$\frac{\mathrm{d}c_1(t)}{\mathrm{d}t} = \frac{i}{2}\Omega\big(e^{i(\omega-\omega_{21})t} + e^{-i(\omega+\omega_{21})t}\big)c_2(t), \qquad (6.37)$$

$$\frac{\mathrm{d}c_2(t)}{\mathrm{d}t} = \frac{i}{2}\Omega\big(e^{i(\omega+\omega_{21})t} + e^{-i(\omega-\omega_{21})t}\big)c_1(t). \qquad (6.38)$$

The above set of equations are in general non-integrable, but can be solved under certain assumptions [59]. First, we assume that the laser field is in resonance with the transition between the two states, which gives $\omega - \omega_{21} = 0$, and thus $e^{\pm i(\omega-\omega_{21})} = 1$. We invoke the rotating wave approximation (RWA), in which the rapidly oscillating terms containing the phase factor $e^{\pm i(\omega+\omega_{21})}$ are neglected. The RWA generally works very well close to resonance [59]. Taking the above assumptions into account, Equation (6.37) and Equation (6.38) simplify to

$$\frac{\mathrm{d}c_1(t)}{\mathrm{d}t} = \frac{i}{2}\Omega c_2(t), \qquad (6.39)$$

$$\frac{\mathrm{d}c_2(t)}{\mathrm{d}t} = \frac{i}{2}\Omega c_1(t). \qquad (6.40)$$

Assuming that the time-dependent state starts out in $|\psi_1\rangle$, which leads to the initial conditions $c_1(t_0) = 1$ and $c_2(t_0) = 0$, these differential equations

have the solutions

$$c_1(t) = \cos\left(\frac{\Omega t}{2}\right), \tag{6.41}$$

$$c_2(t) = i\sin\left(\frac{\Omega t}{2}\right), \tag{6.42}$$

and the time-dependent probabilities of observing the system in the two states are

$$|c_1(t)|^2 = \cos^2\left(\frac{\Omega t}{2}\right) = \frac{1 + \cos(\Omega t)}{2}, \tag{6.43}$$

$$|c_2(t)|^2 = \sin^2\left(\frac{\Omega t}{2}\right) = \frac{1 - \cos(\Omega t)}{2}. \tag{6.44}$$

We can see that the system oscillates sinusoidally between the two states with the Rabi frequency Ω, due to the alternating absorption and emission of radiation. The probability of observing the system in the two states is illustrated as a function of time in Figure 6.3.

Chapter 7

Summary of papers and conclusion

In this chapter, we summarize the five papers that are included in this thesis.

7.1 Paper I: e^T 1.0: An open source electronic structure program with emphasis on coupled cluster and multilevel methods

In the first paper of this thesis, the capabilities of version 1.0 of the electronic structure program e^T are described. The program was developed in collaboration between authors at the Norwegian University of Science and Technology, the Technical University of Denmark and the Scuola Normale Superiore in Italy. The program is written mainly in modern Fortran, and the object-oriented capabilities of the language are used extensively in order to make the code modular and easily extensible. It is inspired by the Dalton program [60], which many of the time-independent methods in e^T were tested against during development.

The e^T code is based on Cholesky decomposed two-electron integrals, leading to reduced storage requirements and increased computational efficiency. Version 1.0 includes well-performing ground state and equation-of-motion coupled-cluster methods based on spin-adapted CCS, CC2, CCSD and CC3, and time-dependent coupled-cluster methods based on spin-adapted CCS and CCSD. See Chapter 5 for an introduction to these methods. The version also includes multilevel methods, where the orbital space is partitioned into active and inactive orbitals and the active orbitals are treated at a higher level of theory than the inactive ones, and

hybrid quantum mechanics/molecular mechanics methods.

A prerequisite for the implementation of time-dependent coupled-cluster methods is the ability to handle complex time-dependent cluster amplitudes and multipliers, which can be seen from Equation (5.48) and Equation (5.49) in Chapter 5. In order to avoid manual maintenance of both the complex code and the more efficient real code used in time-independent methods, we developed scripts that auto-generate most of the complex code from the real one. We also implemented several Runge-Kutta integrators and the capability to calculate interactions with an arbitrary number of pulses, with different linear polarizations, field strengths, carrier frequencies, envelopes and carrier-envelope phases. See Section 5.7 and Section 3.5 for more information about Runge-Kutta methods and the pulse shapes, respectively.

For illustrative purposes, two different TDCCSD simulations of the excitation of a water molecule by linearly polarized Gaussian pulses were performed. The frequency-space dipole moment of the excited molecule was found by Fourier transforming the time-dependent dipole moment, and displayed peaks at the same positions as EOM-CCSD peaks calculated with the Davidson method, which is given a brief introduction in Section 5.4. The intensities of the peaks from the two methods differed, however, due to the finite bandwidth of the pulse used in the TDCC simulation.

7.2   Paper II: Time-dependent coupled-cluster theory for ultrafast transient-absorption spectroscopy

This paper describes the spin-adapted TDCC method implemented in the e^T program in further detail, and includes a demonstration how it can be used for simulating ultrafast transient absorption spectroscopy. More specifically, the method was used to calculate the frequency-dependent response of diatomic molecules to an attosecond valence-exciting pump pulse followed by an attosecond core-exciting probe pulse. Initial simulations were used to ensure the stability of the method with respect to changing model parameters, including the numerical integrator, the basis set and the truncation level of the TDCC method. The results were seen to be reasonably well-behaved with respect to the changing parameters, indicating that the TDCC approach can be used for modeling molecular multiphoton processes that involve bound states targeted by valence- and core-exciting pulses.

We then proceeded to calculate the frequency-resolved transient ab-

sorption, which exhibits variations with the temporal pump-probe separation. In the x-ray frequency region, the peaks of the transient absorption spectrum were seen to oscillate as a linear combination of a few sinusoidal functions. This suggests that the TDCC method can model the phenomenon of quantum beats occurring in x-ray transient absorption spectroscopy, see Section 6.3. We also observed that the oscillations and positions of the peaks in the spectrum can be interpreted in terms of energy differences of the different field-free EOM-CC eigenstates involved in the laser-molecule interactions.

7.3 Paper III: Simulating weak-field attosecond processes with a Lanczos reduced basis approach to time-dependent equation-of-motion coupled-cluster theory

This paper describes the generation of field-free valence- and core-excited EOM-CC states with an asymmetric band Lanczos algorithm, and the use of these states in simulations of attosecond pump-probe spectroscopy with the TD-EOM-CC method. Several extensions to the e^T program were made for this paper. First, the asymmetric band Lanczos algorithm in Ref. [61] was implemented, where we modified the biorthogonalization steps in this algorithm for improved numerical stability. See Section 5.4 for a brief introduction to Lanczos methods. We furthermore implemented starting vectors based on excited EOM-CC states, and a Rayleigh-Ritz procedure for combining EOM-CC states from different calculations. The program was modified to handle the integration of the TD-EOM-CC equations in the basis of field-free EOM-CC eigenstates, see Section 5.6. Embedded integrators with adaptive time steps were also implemented, for increased stability and ease of integration as described in Section 5.7.

The particular starting vectors direct the asymmetric band Lanczos algorithm toward field-free EOM-CC states contributing to the interaction with pump and probe pulses in weak-field transient absorption spectroscopy. In the basis of a few such states, we demonstrated that TD-EOM-CC largely reproduces the TDCC results in Paper II, and the removal of irrelevant degrees of freedom makes the approach significantly more computationally efficient than TDCC. This suggests that the use of TD-EOM-CC can be warranted in the weakly non-linear regime. From additional simulations of the ethylene molecule in the paper, however, the results also indicate that TD-EOM-CC can behave differently from

TDCC for certain systems and pulses, but this claim was not investigated further in this paper. We furthermore investigated the effect of the CVS approximation, see Section 5.3. The approximation was seen to work well for the *K*-edges of heavy nuclei, but for lighter nuclei we noticed the presence of valence pseudo-continuum states with energies around the *K*-edge, complicating the simulation and interpretation of the results. The paper also includes a TD-EOM-CC simulation of the integrated transient absorption of glycine at the oxygen *K*-edge. We noted that the absorption correlates with the projection of the dipole moment in the direction from the molecular center-of-mass to the oxygen atom, suggesting that the probe pulse reveals the charge migration in this direction.

7.4   Paper IV: Comparing real-time coupled cluster methods through simulation of collective Rabi oscillations

For this paper, the e^T program was extended for handling TD-EOM-CC methods in the determinant basis, see Section 5.6, which is equivalent to expressing TD-EOM-CC in terms of all field-free EOM-CC states.

   Theoretically, we demonstrated that the TDCC method gives properties that scale correctly with respect to system size, irrespective of the strength of external field, and also that this is not the case for the TD-EOM-CC method unless there is no interaction with the field. The scaling properties of the two methods were further assessed through simulations of distant atoms interacting with an electromagnetic field chosen to be resonant with transitions of helium atoms in the simulations. This interaction gives rise to Rabi oscillations, see Section 6.4. In exact quantum theory, the frequency of Rabi oscillations should not scale with the number of identical and non-interacting systems in resonance with the field. From TD-EOM-CCSD simulations, however, the Rabi frequency was shown to increase as the square root of number of subsystems in resonance with the field, but also to not increase with the number of non-resonant subsystems. This suggests that TD-EOM-CCSD does not give a qualitatively correct description of collective Rabi oscillations, but can be used for simulating solitary Rabi oscillations in extended systems.

   We further expanded on the observation that TDCCSD fails for reference depletion made in Ref. [17] by showing that simulations also fail when the system of multiple distant atoms approaches complete population inversion. We observed two reasons for why this happens. First, the cluster amplitudes blow up due to the two-center Coulomb integrals, which

are small in magnitude, not being exactly equal to zero. Second, the left amplitudes that do not contribute to time-dependent expectation values for non-interacting systems also blow up. Nevertheless, the TDCC method still exhibits the correct scaling properties up to the point of instabilities, and we demonstrated that the simulations of multiple distant atoms can be made to finish by eliminating the two-center Coulomb integrals and the badly behaved left amplitudes. This modification did not affect the energy expectation value significantly, except for removing a slight blowup of the value occurring at the point of failure. After these modifications were made to the TDCC method, the scaling properties of the Rabi frequencies were shown to be correct with respect to the numbers of resonant and off-resonant atoms.

7.5 Paper V: Coupled cluster simulation of impulsive stimulated x-ray Raman scattering

In this paper, the diagonal basis TD-EOM-CCSD method was used to simulate impulsive x-ray Raman scattering by molecules, where an attosecond x-ray pulse transfers a fraction of the electronic population of molecules to valence-excited states, through a Raman process going through intermediate core-excited states, as described by the second-order term of the Dyson series in Section 6.1. The Raman process was tracked by calculating the time-dependent and final populations of the relevant field-free EOM-CC states, and by calculating and visualizing the difference between the electronic time-dependent and ground state densities. We demonstrated that the relative orientation of the molecule and laser pulses has great importance for the yield of the Raman process for symmetry reasons, see Section 6.2. The time-dependent electronic density difference was furthermore used to demonstrate that the intermediate core excitations of the Raman process are local to the atom with the K-edge targeted by the attosecond pulse, and that the valence-excited wavepacket is created in approximately a single femtosecond. This supports the claim that attosecond x-ray pulses can be used for initiating photochemical processes with high temporal and spatial precision.

7.6 Summary and outlook

The TDCC and TD-EOM-CC approaches described in this thesis enabled us to simulate electronic multiphoton processes in atoms and molecules in

the real-time domain. The details of the approaches and simulations were reported in a total of five papers, including explorations and comparisons of the properties of the TDCC and TD-EOM-CC methods. In Paper II and Paper III, we demonstrated how the TDCC and TD-EOM-CC methods can be used for modeling attosecond transient absorption spectroscopy. Furthermore, in Paper III we observed that the two methods can give similar results for small systems in weak fields. In Paper IV, on the other hand, we demonstrated that the scaling properties of TDCC agree with the properties of exact quantum theory, while the scaling properties of TD-EOM-CC can be incorrect, in particular for systems with multiple subsystems in resonance with a strong field. The TDCC method is, however, seen to fail when certain multi-atom systems approach complete population inversion. The TD-EOM-CC method is easier to handle: the method does not exhibit the instabilities seen for TDCC in the simulations of Paper IV, and we demonstrate how a reduced-basis formulation can be used to increase the numerical efficiency of the method in Paper III. In Paper V we also show how the time-dependent TD-EOM-CC state can be interpreted in terms of eigenstate probabilities, while it is not clear how this can be done consistently for TDCC [62]. In conclusion, the methods have their own sets of strengths and weaknesses, and the method should be chosen with regard to the nature of the physical process of interest.

The approaches considered in this thesis are far from giving a complete picture of the physical processes occurring in the experiments we have attempted to simulate. The nuclear degrees of freedom have been neglected. So has the coupling of the atoms and molecules to their environment, including the coupling to the vacuum electromagnetic field, which causes spontaneous emission. We have assumed fixed relative orientations between the atom or molecule and the polarization of the electromagnetic field, and also precise control over the phases and amplitudes of attosecond laser pulses. This is often not realized in experiments. Furthermore, spin and relativistic effects have been neglected, and we have assumed the validity of the dipole approximation, which breaks down in the limit of large systems and strong fields. The employed Gaussian basis sets are very sparse, and must be amended or replaced before Rydberg states and ionization can be described properly. Many of the shortcomings of the approaches can be addressed, but remedies are often associated with an increase in computational cost. In Paper II and Paper III we did, however, target the lowest-energy weak-field valence- and core-excitations of moderately sized molecules with light nuclei, shortly after excitation, which should be reasonably well represented within the chosen approximations. Also, Paper IV focuses more on the internal workings of the implemented methods, which

makes the faithful representation of the strong laser-molecule interactions less important. Importantly, we have demonstrated that truncated real-time coupled-cluster methods have several properties that are consistent with the behavior of untruncated methods, which motivates their further use for simulating molecular quantum dynamics.

Bibliography

(1) Poltev, V. In *Handbook of Computational Chemistry*, Leszczynski, J., Ed.; Springer Netherlands: Dordrecht, 2016, pp 1–48.

(2) Van Gunsteren, W. F.; Bakowies, D.; Baron, R.; Chandrasekhar, I.; Christen, M.; Daura, X.; Gee, P.; Geerke, D. P.; Glättli, A.; Hünenberger, P. H.; Kastenholz, M. A.; Oostenbrink, C.; Schenk, M.; Trzesniak, D.; van der Vegt, N. F. A.; Yu, H. B. Biomolecular Modeling: Goals, Problems, Perspectives. *Angewandte Chemie International Edition* **2006**, *45*, 4064–4092.

(3) Hofer, T. S.; de Visser, S. P. Editorial: Quantum Mechanical/Molecular Mechanical Approaches for the Investigation of Chemical Systems – Recent Developments and Advanced Applications. *Frontiers in Chemistry* **2018**, *6*, DOI: `10.3389/fchem.2018.00357`.

(4) Whitfield, J. D.; Love, P. J.; Aspuru-Guzik, A. Computational complexity in electronic structure. *Physical Chemistry Chemical Physics* **2013**, *15*, 397–411.

(5) Ivanov, M. Concluding remarks: The age of molecular movies. *Faraday Discussions* **2021**, *228*, 622–629.

(6) Solá, I. R.; González-Vázquez, J.; de Nalda, R.; Bañares, L. Strong field laser control of photochemistry. *Physical Chemistry Chemical Physics* **2015**, *17*, 13183–13200.

(7) Huang, N.; Deng, H.; Liu, B.; Wang, D.; Zhao, Z. Features and futures of X-ray free-electron lasers. *The Innovation* **2021**, *2*, 100097.

(8) Midorikawa, K. Progress on table-top isolated attosecond light sources. *Nature Photonics* **2022**, *16*, 267–278.

(9) Dutoi, A. D.; Gokhberg, K.; Cederbaum, L. S. Time-resolved pump-probe spectroscopy to follow valence electronic motion in molecules: Theory. *Physical Review A* **2013**, *88*, 013419.

(10) Marangos, J. P. Accessing the quantum spatial and temporal scales with XFELs. *Nature Reviews Physics* **2020**, *2*, 332–334.

(11) Bishop, R. F. An overview of coupled cluster theory and its applications in physics. *Theoretica chimica acta* **1991**, *80*, 95–148.

(12) Bartlett, R. J.; Musia, M. Coupled-cluster theory in quantum chemistry. *Reviews of Modern Physics* **2007**, *79*, 291–352.

(13) Li, X.; Govind, N.; Isborn, C.; DePrince, A. E. I.; Lopata, K. Real-Time Time-Dependent Electronic Structure Theory. *Chemical Reviews* **2020**, *120*, PMID: 32813506, 9951–9993.

(14) Kvaal, S. Ab initio quantum dynamics using coupled-cluster. *The Journal of Chemical Physics* **2012**, *136*, 194109.

(15) Sonk, J. A.; Caricato, M.; Schlegel, H. B. TD-CI Simulation of the Electronic Optical Response of Molecules in Intense Fields: Comparison of RPA, CIS, CIS(D), and EOM-CCSD. *The Journal of Physical Chemistry A* **2011**, *115*, PMID: 21495664, 4678–4690.

(16) Luppi, E.; Head-Gordon, M. Computation of high-harmonic generation spectra of H2 and N2 in intense laser pulses using quantum chemistry methods and time-dependent density functional theory. *Molecular Physics* **2012**, *110*, 909–923.

(17) Pedersen, T. B.; Kvaal, S. Symplectic integration and physical interpretation of time-dependent coupled-cluster theory. *The Journal of Chemical Physics* **2019**, *150*, 144106.

(18) Kristiansen, H. E.; Schøyen, Ø. S.; Kvaal, S.; Pedersen, T. B. Numerical stability of time-dependent coupled-cluster methods for many-electron dynamics in intense laser pulses. *The Journal of Chemical Physics* **2020**, *152*, 071102, DOI: 10.1063/1.5142276.

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

(20) Stefanucci, G.; van Leeuwen, R., *Nonequilibrium Many-Body Theory of Quantum Systems: A Modern Introduction*; Cambridge University Press: 2013.

(21) In *Theoretical Femtosecond Physics: Atoms and Molecules in Strong Laser Fields*; Springer Berlin Heidelberg: Berlin, Heidelberg, 2008, pp 77–95.

(22) In *Molecular Physics*; John Wiley & Sons, Ltd: 2005; Chapter 2, pp 15–77.

(23) Cederbaum, L. S. BornOppenheimer approximation and beyond for time-dependent electronic processes. *The Journal of Chemical Physics* **2008**, *128*, 124101.

(24) Griffiths, D. J., *Introduction to Electrodynamics*, 4th ed.; Cambridge University Press: 2017.

(25) In *Methods of Mathematical Physics*; John Wiley & Sons, Ltd: 1989; Chapter 3, pp 154–239.

(26) Maurer, J.; Keller, U. Ionization in intense laser fields beyond the electric dipole approximation: concepts, methods, achievements and future directions. *Journal of Physics B: Atomic, Molecular and Optical Physics* **2021**, *54*, 094001.

(27) 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.

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

(29) Helgaker, T.; Jørgensen, P.; Olsen, J. In *Molecular ElectronicStructure Theory*; John Wiley & Sons, Ltd: 2000; Chapter 1, pp 1–33.

(30) Roothaan, C. C. J. New Developments in Molecular Orbital Theory. *Reviews of Modern Physics* **1951**, *23*, 69–89.

(31) Beebe, N. H. F.; Linderberg, J. Simplifications in the generation and transformation of two-electron integrals in molecular calculations. *International Journal of Quantum Chemistry* **1977**, *12*, 683–705.

(32) Koch, H.; Sánchez de Merás, A.; Pedersen, T. B. Reduced scaling in electronic structure calculations using Cholesky decompositions. *The Journal of Chemical Physics* **2003**, *118*, 9481–9484.

(33) Folkestad, S. D.; Kjønstad, E. F.; Myhre, R. H.; Andersen, J. H.; Balbi, A.; Coriani, S.; Giovannini, T.; Goletto, L.; Haugland, T. S.; Hutcheson, A.; Høyvik, I.-M.; Moitra, T.; Paul, A. C.; Scavino, M.; Skeidsvoll, A. S.; Tveten, Å. H.; Koch, H. eT 1.0: An open source electronic structure program with emphasis on coupled cluster and multilevel methods. *The Journal of Chemical Physics* **2020**, *152*, 184103.

(34) Hall, G. G.; Lennard-Jones, J. E. 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.

(35) Koch, H.; Kobayashi, R.; Sanchez de Merás, A.; Jørgensen, P. Calculation of sizeintensive transition moments from the coupled cluster singles and doubles linear response function. *The Journal of Chemical Physics* **1994**, *100*, 4393–4400.

(36) íek, J. On the Correlation Problem in Atomic and Molecular Systems. Calculation of Wavefunction Components in UrsellType Expansion Using QuantumField Theoretical Methods. *The Journal of Chemical Physics* **1966**, *45*, 4256–4266.

(37) Stanton, J. F.; Bartlett, R. J. The equation of motion coupledcluster method. A systematic biorthogonal approach to molecular excitation energies, transition probabilities, and excited state properties. *The Journal of Chemical Physics* **1993**, *98*, 7029–7039.

(38) 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.

(39) 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.

(40) Lanczos, C. An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *Journal of research of the National Bureau of Standards* **1950**, *45*, 255–282.

(41) Chen, T.; Demmel, J.; Gu, M.; Saad, Y.; Lehoucq, R.; Sorensen, D.; Maschhoff, K.; Bai, Z.; Day, D.; Freund, R.; Sleijpen, G.; van der Vorst, H.; Li, R. In *Templates for the Solution of Algebraic Eigenvalue Problems*, pp 149–231.

(42) Morgan, R. B.; Scott, D. S. Generalizations of Davidsons Method for Computing Eigenvalues of Sparse Symmetric Matrices. *SIAM Journal on Scientific and Statistical Computing* **1986**, *7*, 817–825.

(43) Koch, H.; Jörgensen, P. Coupled cluster response functions. *The Journal of Chemical Physics* **1990**, *93*, 3333–3344.

(44) Helgaker, T.; Coriani, S.; Jørgensen, P.; Kristensen, K.; Olsen, J.; Ruud, K. Recent Advances in Wave Function-Based Methods of Molecular-Property Calculations. *Chemical Reviews* **2012**, *112*, PMID: 22236047, 543–631.

(45) Nascimento, D. R.; DePrince, A. E. I. Linear Absorption Spectra from Explicitly Time-Dependent Equation-of-Motion Coupled-Cluster Theory. *Journal of Chemical Theory and Computation* **2016**, *12*, 5834–5840.

(46) Nascimento, D. R.; DePrince, A. E. I. Simulation of Near-Edge X-ray Absorption Fine Structure with Time-Dependent Equation-of-Motion Coupled-Cluster Theory. *The Journal of Physical Chemistry Letters* **2017**, *8*, 2951–2957.

(47)   Nascimento, D. R.; DePrince A. Eugene, I. A general time-domain formulation of equation-of-motion coupled-cluster theory for linear spectroscopy. *The Journal of Chemical Physics* **2019**, *151*, DOI: 10.1063/1.5125494.

(48)   Park, Y. C.; Perera, A.; Bartlett, R. J. Equation of motion coupled-cluster for core excitation spectra: Two complementary approaches. *The Journal of Chemical Physics* **2019**, *151*, DOI: 10.1063/1.5117841.

(49)   Iserles, A., *A First Course in the Numerical Analysis of Differential Equations*, 2nd ed.; Cambridge Texts in Applied Mathematics; Cambridge University Press: 2008.

(50)   Dyson, F. J. The Radiation Theories of Tomonaga, Schwinger, and Feynman. *Physical Review* **1949**, *75*, 486–502.

(51)   Cao, J.; Wilson, K. R. Detecting wave packet motion in pumpprobe experiments: Theoretical analysis. *The Journal of Chemical Physics* **1997**, *106*, 5062–5072.

(52)   Uberna, R.; Khalil, M.; Williams, R. M.; Papanikolas, J. M.; Leone, S. R. Phase and amplitude control in the formation and detection of rotational wave packets in the E1g+ state of Li2. *The Journal of Chemical Physics* **1998**, *108*, 9259–9274.

(53)   Ballard, J. B.; Stauffer, H. U.; Amitay, Z.; Leone, S. R. Optimization of wave packet coefficients in Li2 using an evolutionary algorithm: The role of resonant and nonresonant wavelengths. *The Journal of Chemical Physics* **2002**, *116*, 1350–1360.

(54)   Ballard, J. B.; Dai, X.; Arrowsmith, A. N.; Hüwel, L.; Stauffer, H. U.; Leone, S. R. Observation of wave packets with simultaneous electronic, vibrational, and rotational degrees of freedom in Li2. *Chemical Physics Letters* **2005**, *402*, 27–31.

(55)   Palacios, A.; González-Castrillo, A.; Martín, F. Molecular interferometer to decode attosecond electron–nuclear dynamics. *Proceedings of the National Academy of Sciences* **2014**, *111*, 3973–3978.

(56)   Bencivenga, F.; Capotondi, F.; Principi, E.; Kiskinova, M.; Masciovecchio, C. Coherent and transient states studied with extreme ultraviolet and X-ray free electron lasers: present and future prospects. *Advances in Physics* **2014**, *63*, 327–404.

(57)   Li, K.; Laksman, J.; Mazza, T.; Doumy, G.; Koulentianos, D.; Picchiotti, A.; Serkez, S.; Rohringer, N.; Ilchen, M.; Meyer, M.; Young, L. Ghost-imaging-enhanced noninvasive spectral characterization of stochastic x-ray free-electron-laser pulses. *Communications Physics* **2022**, *5*, 191.

(58)   Driver, T.; Li, S.; Champenois, E. G.; Duris, J.; Ratner, D.; Lane, T. J.; Rosenberger, P.; Al-Haddad, A.; Averbukh, V.; Barnard, T.; Berrah, N.; Bostedt, C.; Bucksbaum, P. H.; Coffee, R.; DiMauro, L. F.; Fang, L.; Garratt, D.; Gatton, A.; Guo, Z.; Hartmann, G.; Haxton, D.; Helml, W.; Huang, Z.; LaForge, A.; Kamalov, A.; Kling, M. F.; Knurr, J.; Lin, M.-F.; Lutman, A. A.; MacArthur, J. P.; Marangos, J. P.; Nantel, M.; Natan, A.; Obaid, R.; O'Neal, J. T.; Shivaram, N. H.; Schori, A.; Walter, P.; Li Wang, A.; Wolf, T. J. A.; Marinelli, A.; Cryan, J. P. Attosecond transient absorption spooktroscopy: a ghost imaging approach to ultrafast absorption spectroscopy. *Physical Chemistry Chemical Physics* **2020**, *22*, 2704–2712.

(59)   Liu, J.; Li, Z.-Y. Interaction of a two-level atom with single-mode optical field beyond the rotating wave approximation. *Optics Express* **2014**, *22*, 28671–28682.

(60)   Aidas, K.; Angeli, C.; Bak, K. L.; Bakken, V.; Bast, R.; Boman, L.; Christiansen, O.; Cimiraglia, R.; Coriani, S.; Dahle, P.; Dalskov, E. K.; Ekström, U.; Enevoldsen, T.; Eriksen, J. J.; Ettenhuber, P.; Fernández, B.; Ferrighi, L.; Fliegl, H.; Frediani, L.; Hald, K.; Halkier, A.; Hättig, C.; Heiberg, H.; Helgaker, T.; Hennum, A. C.; Hettema, H.; Hjertenæs, E.; Høst, S.; Høyvik, I.-M.; Iozzi, M. F.; Jansík, B.; Jensen, H. J. A.; Jonsson, D.; Jørgensen, P.; Kauczor, J.; Kirpekar, S.; Kjærgaard, T.; Klopper, W.; Knecht, S.; Kobayashi, R.; Koch, H.; Kongsted, J.; Krapp, A.; Kristensen, K.; Ligabue, A.; Lutnæs, O. B.; Melo, J. I.; Mikkelsen, K. V.; Myhre, R. H.; Neiss, C.; Nielsen, C. B.; Norman, P.; Olsen, J.; Olsen, J. M. H.; Osted, A.; Packer, M. J.; Pawlowski, F.; Pedersen, T. B.; Provasi, P. F.; Reine, S.; Rinkevicius, Z.; Ruden, T. A.; Ruud, K.; Rybkin, V. V.; Saek, P.; Samson, C. C. M.; de Merás, A. S.; Saue, T.; Sauer, S. P. A.; Schimmelpfennig, B.; Sneskov, K.; Steindal, A. H.; Sylvester-Hvid, K. O.; Taylor, P. R.; Teale, A. M.; Tellgren, E. I.; Tew, D. P.; Thorvaldsen, A. J.; Thøgersen, L.; Vahtras, O.; Watson, M. A.; Wilson, D. J. D.; Ziolkowski, M.; Ågren, H. The Dalton quantum chemistry program system. *WIREs Computational Molecular Science* **2014**, *4*, 269–284.

(61)   Freund, R. W. Model reduction methods based on Krylov subspaces. *Acta Numerica* **2003**, *12*, 267–319.

(62) Pedersen, T. B.; Kristiansen, H. E.; Bodenstein, T.; Kvaal, S.; Schøyen, Ø. S. Interpretation of Coupled-Cluster Many-Electron Dynamics in Terms of Stationary States. *Journal of Chemical Theory and Computation* **2021**, *17*, PMID: 33337895, 388–404.

Paper I

e^T 1.0: An open source electronic structure program with emphasis on coupled cluster and multilevel methods

Sarai D. Folkestad,[1] Eirik F. Kjønstad,[1] Rolf H. Myhre,[1] Josefine H. Andersen,[2] Alice Balbi,[3] Sonia Coriani,[2] Tommaso Giovannini,[1] Linda Goletto,[1] Tor S. Haugland,[1] Anders Hutcheson,[1] Ida-Marie Høyvik,[1] Torsha Moitra,[2] Alexander C. Paul,[1] Marco Scavino,[3] Andreas S. Skeidsvoll,[1] Åsmund H. Tveten,[1] and Henrik Koch[1,3,a]

AFFILIATIONS

[1] Department of Chemistry, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway
[2] DTU Chemistry—Department of Chemistry, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark
[3] Scuola Normale Superiore, Piazza dei Cavalieri, 7, IT-56126 Pisa, PI, Italy

Note: This article is part of the JCP Special Topic on Electronic Structure Software.
[a] Author to whom correspondence should be addressed: henrik.koch@sns.it

ABSTRACT

The e^T program is an open source electronic structure package with emphasis on coupled cluster and multilevel methods. It includes efficient spin adapted implementations of ground and excited singlet states, as well as equation of motion oscillator strengths, for CCS, CC2, CCSD, and CC3. Furthermore, e^T provides unique capabilities such as multilevel Hartree–Fock and multilevel CC2, real-time propagation for CCS and CCSD, and efficient CC3 oscillator strengths. With a coupled cluster code based on an efficient Cholesky decomposition algorithm for the electronic repulsion integrals, e^T has similar advantages as codes using density fitting, but with strict error control. Here, we present the main features of the program and demonstrate its performance through example calculations. Because of its availability, performance, and unique capabilities, we expect e^T to become a valuable resource to the electronic structure community.

Published under license by AIP Publishing. https://doi.org/10.1063/5.0004713

I. INTRODUCTION

During the last five decades, a wide variety of models and algorithms have been developed within the field of electronic structure theory and many program packages are now available to the community.[1] Programs with extensive coupled cluster functionality include CFOUR,[2] Dalton,[3] GAMESS,[4] Gaussian,[5] Molcas,[6] Molpro,[7] NWChem,[8] ORCA,[9] PSI4,[10] QChem,[11] and TURBOMOLE.[12] Although these are all general purpose quantum chemistry programs, each code is particularly feature rich or efficient in specific areas. For instance, a large variety of response properties[13] have been implemented in Dalton, CFOUR is particularly suited for gradients[14,15] and geometry optimization, and QChem is leading in equation of motion[16,17] (EOM) features. However, due to the long

history of many of these programs, it can be challenging to modify and optimize the existing features or to integrate new methods and algorithms.

In 2016, we began developing a coupled cluster code based on Cholesky decomposed electron repulsion integrals.[18,19] While starting anew, we have drawn inspiration from Dalton[3] and used it extensively for testing purposes. Our goal is to create an efficient, flexible, and easily extendable foundation upon which coupled cluster methods and features—both established and new—can be developed. That code has now evolved beyond a coupled cluster code into a freestanding electronic structure program. It is named e^T after the expression for the coupled cluster ground state wave function,[20]

$$|\Psi\rangle = e^T |\mathrm{R}\rangle, \tag{1}$$

and released as an open source program licensed under the GNU General Public License 3 (GPL 3.0).

The first version of e^T offers an optimized Hartree–Fock (HF) code and a wide range of standard coupled cluster methods. It includes the most efficient published implementations of Cholesky decomposition of the electron repulsion integrals[21] and of coupled cluster singles, doubles, and perturbative triples[22,23] (CC3). Furthermore, e^T features the first released implementations of multilevel HF[24] (MLHF), multilevel coupled cluster singles and perturbative doubles[25,26] (MLCC2), and explicitly time-dependent coupled cluster singles (TD-CCS), and singles and doubles (TD-CCSD) theory. All coupled cluster models can be used in quantum mechanics/molecular mechanics[27,28] (QM/MM) calculations or be combined with the polarizable continuum model[29,30] (PCM).

e^T is primarily written in modern Fortran using the Fortran 2008 standard. The current version of the code is interfaced to two external libraries: Libint 2[31] for the atomic orbital integrals and PCMSolver 1.2[32] for PCM embedding. In addition, e^T applies the runtest library[33] for testing and a CMake module from autocmake[34] to locate and configure BLAS and LAPACK.

With the introduction of the 2003 and 2008 standards, Fortran has become an object oriented programming language. We have exploited this to make e^T modular, readable, and easy to extend. Throughout the program, we use OpenMP[35] to parallelize computationally intensive loops and BLAS and LAPACK routines wherever possible. In order to preserve code quality, extensive code review and enforcement of a consistent standard have been prioritized from the outset. While this requires extra effort from both developers and maintainers, it pays dividends in code readability and flexibility.

II. PROGRAM FEATURES

A. Coupled cluster methods

The e^T program features all standard coupled cluster methods up to perturbative triples: singles (CCS), singles with perturbative doubles[36] (CC2), singles and doubles[37] (CCSD), singles and doubles with non-iterative perturbative triples[38] [CCSD(T)], and singles and doubles with perturbative triples[22] (CC3). At the CCSD(T) level of theory, only ground state energies can be computed. For all other methods, efficient spin adapted implementations of ground and excited singlet states are available. Moreover, dipole and quadrupole moments, as well as EOM oscillator strengths, can be calculated. Equation of motion polarizabilities are available at the CCS, CC2, and CCSD levels of theory.

A number of algorithms are implemented to solve the coupled cluster equations. For linear and eigenvalue equations, we have implemented the Davidson method.[39] This algorithm is used to solve the ground state multiplier equations, response equations, and excited state equations. To handle nonlinear coupled cluster equations, we have implemented algorithms that use direct inversion of the iterative subspace[40,41] (DIIS) to accelerate convergence. The ground state amplitude equations can be solved using DIIS combined with the standard[1,42] quasi-Newton algorithm or exact Newton–Raphson. We also use a DIIS-accelerated algorithm[43] for

the nonlinear excited state equations in CC2 and CC3. Our implementation of DIIS incorporates the option to use the related conjugate residual with optimal trial vectors[44,45] (CROP) method for acceleration. For the nonperturbative coupled cluster methods, the asymmetric Lanczos algorithm is also available.[46,47]

The time-dependent coupled cluster equations can be explicitly solved for CCS and CCSD[48,49] using Euler, Runge–Kutta 4 (RK4), or Gauss–Legendre (GL2, GL4, and GL6) integrators. This requires implementations of the amplitude and multiplier equations with complex variables. Any number of classical electromagnetic pulses can be specified in the length gauge, assuming that the dipole approximation is valid. A modified version of the fast Fourier transform library FFTPACK 5.1[50] is used to extract frequency domain information.

B. Cholesky decomposition for the electronic repulsion integrals

Cholesky decomposition is an efficient method to obtain a compact factorization of the rank deficient electron repulsion integral matrix.[18,19,51] All post-HF methods in e^T rely on the Cholesky vectors to construct the electron repulsion integrals. One advantage of factorization is the reduced storage requirements; the size of the Cholesky vectors scales as $\mathcal{O}(n_{AO}^3)$, while the full integral matrix scales as $\mathcal{O}(n_{AO}^4)$. The Cholesky vectors are kept in memory when possible but are otherwise stored on disk. Another advantage is that they allow for an efficient construction and transformation of subsets of the integrals. The Cholesky decomposition in e^T is highly efficient, consisting of a two-step procedure that reduces both storage requirements and computational cost compared to earlier algorithms. For a description of the algorithm and performance comparisons to Molcas,[6] see Ref. 21.

C. Hartree–Fock

The restricted HF (RHF) and unrestricted HF (UHF) models are implemented in e^T. The implementations are integral direct and exploit Coloumb and exchange screening and permutation symmetry. We use a superposition of atomic densities[52] (SAD) initial guess constructed from spherically averaged UHF calculations on the constituent atoms. The Hartree–Fock equations are solved using a Roothan–Hall self-consistent field (SCF) algorithm accelerated by either DIIS or CROP. To improve the screening and reduce the number of integrals that must be evaluated, density differences are used to construct the Fock matrix.

D. Multilevel and multiscale methods

In MLHF, a region of the molecular system is defined as active. A set of active occupied orbitals are obtained through a restricted, partial Cholesky decomposition of an initial idempotent AO density matrix.[53] The active virtual orbitals are obtained by constructing projected atomic orbitals[54,55] (PAOs) centered on the active atoms. The PAOs are orthonormalized through the canonical orthonormalization procedure.[56] The MLHF equations are solved using a DIIS accelerated, MO based, Roothan–Hall SCF algorithm. Only the active MOs are optimized.[57]

The most expensive step of an MLHF calculation is the construction of the inactive two-electron contribution to the Fock

matrix. As the inactive orbitals are frozen, it is only necessary to calculate this term once. The iterative cost in MLHF is dominated by the construction of the active two-electron contribution to the Fock matrix. An additional Coulomb and exchange screening, which targets accuracy of the matrix in the active MO basis, reduces the cost. The active orbitals are localized, and consequently, the elements of the AO Fock matrix that correspond to AOs distant from the active atoms will not significantly contribute to the active MO Fock matrix. This is similar to the screening used in MLHF specific Cholesky decomposition of the electron repulsion integrals.[21]

In MLCC2,[23,25,26,58] an active orbital space is treated at the CC2 level of theory, while the remaining inactive orbitals are treated at the CCS level of theory. MLCC2 excitation energies are implemented in e^T. The active space is constructed using the approximated correlated natural transition orbitals,[59,60] Cholesky orbitals, or Cholesky occupied orbitals and PAOs spanning the virtual space.

Frozen orbitals are implemented for all coupled cluster methods in e^T. In addition to the standard frozen core (FC) approximation, reduced space coupled cluster calculations can be performed using semi-localized orbitals. This type of calculation is suited to describe localized properties. In reduced space calculations, the occupied space is constructed from Cholesky orbitals, and PAOs are used to generate the virtual space.

Two QM/MM approaches are available in e^T: electrostatic QM/MM embedding[61] and the polarizable QM/Fluctuating Charge[62] (QM/FQ) model. In the former, the QM density interacts with a set of fixed charges placed in the MM part of the system.[61] In QM/FQ, the QM and MM parts mutually polarize. Each atom in the MM part has a charge that varies as a response to differences in atomic electronegativities and the QM potential.[62] These charges enter the QM Hamiltonian through a term that is nonlinear in the QM density.[63]

PCM embedding can be used in e^T for an implicit description of the external environment. A solute is described at the QM level and is placed in a molecule shaped cavity. The environment is described in terms of an infinite, homogeneous, continuum dielectric that mutually polarizes with the QM part, as in QM/FQ.[64]

In the QM/PCM and QM/FQ implementations, additional terms are only added to the Fock matrix. Additional terms at the coupled cluster level can also be considered.[65–69]

E. Spectroscopic properties and response methods

Coupled cluster is one of the most accurate methods for modeling spectroscopic properties, and both ultraviolet-visible (UV/vis) and x-ray absorption spectra can be modeled in e^T. Core excitations are obtained through the core valence separation (CVS) approximation.[70] CVS is implemented as a projection[71,72] for CCS, CC2, MLCC2, and CCSD. For CC3, amplitudes and excitation vector elements that do not contribute are not calculated. This reduces the scaling of the iterative computational cost for excited states from $\mathcal{O}(n_{MO}^7)$ to $\mathcal{O}(n_{MO}^6)$.

Intensities are obtained from EOM oscillator strengths,[16,17] which are available for CCS, CC2, CCSD, and CC3. In addition, linear response[48] (LR) oscillator strengths can be calculated at the CCS level of theory. The asymmetric Lanczos algorithm[46,47] can be used to directly obtain both energies and EOM oscillator strengths

for CCS, CC2, and CCSD. It can also be combined with the CVS approximation.

Real-time propagation offers a nonperturbative approach to model absorption spectra. Following an initial pulse that excites the system, the dipole moment from the subsequent time evolution can be Fourier transformed to extract the excitation energies and intensities.

Valence ionization potentials are implemented for CCS, CC2, and CCSD. A bath orbital that does not interact with the system is added to the calculation. Excitation vector components not involving this orbital are projected out in an approach similar to the projection in CVS.[71,72]

III. ILLUSTRATIVE APPLICATIONS AND PERFORMANCE TESTS

In this section, we will demonstrate some of the capabilities of e^T with example calculations. Energy thresholds refer to the change in energy from the previous iteration. The maximum norm of the gradient vector is used in Hartree–Fock calculations. For coupled cluster calculations in e^T and Dalton, residual thresholds refer to the L^2 norm of the residual vectors. Finally, the Cholesky decomposition threshold refers to the largest absolute error on the diagonal of the electron repulsion integral matrix. This threshold gives an upper bound to the error of all matrix elements. Coupled cluster calculations were performed with either Cholesky vectors or electron repulsion integrals in memory. All geometries are available from Ref. 73.

A. Coupled cluster methods

The CC2 method is known to yield excitation energies with errors of about 0.1–0.4 eV for valence states with single excitation character.[74–76] The iterative cost of CC2 scales as $\mathcal{O}(n_{MO}^5)$, and it may be implemented with an $\mathcal{O}(n_{MO}^2)$ memory requirement. In Table I, we report the lowest FC-CC2/aug-cc-pVDZ excitation energy of the antibiotic rifampicin[77] (chemical formula $C_{43}H_{58}N_4O_{12}$, see Fig. 1). The calculated excitation energy is 2.58 eV, which is consistent with the orange color of the compound. The ground state was converged to a residual threshold of 10^{-6}, and the excited state was converged to residual and energy thresholds of 10^{-3} and 10^{-6}, respectively. We used a Cholesky decomposition threshold of 10^{-2}, which is sufficient to ensure accuracy of excitation energies in CC2 and CCSD (see Table IV). The calculation was performed on two Intel Xeon Gold 6138 processors using 40 threads and 360 GB shared memory. The average iteration time for the ground state equations was 73 min, and the average iteration time for the excited state equations was 9 h.

TABLE I. The lowest FC-CC2/aug-cc-pVDZ excitation energy (ω) of rifampicin. n_{frozen} is the number of frozen core orbitals.

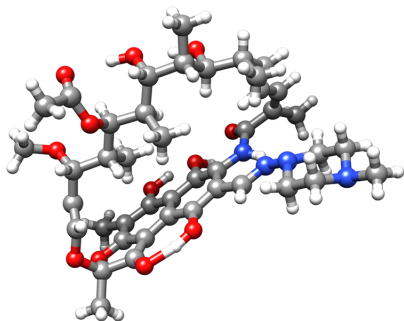| n_{AO}	n_{MO}	n_{frozen}	ω
1879	1865	59	2.579 eV

FIG. 1. Rifampicin.

At the CCSD level of theory, we report calculations for the amino acid tryptophan[78] (chemical formula $C_{11}H_{12}N_2O_2$) and excitation energies for the psychoactive agent lysergic acid diethylamide (LSD)[79] (chemical formula $C_{20}H_{25}N_3O$). Tryptophan and LSD are depicted in Fig. 2.

For tryptophan, we have determined the four lowest excitation energies and the corresponding oscillator strengths at the CCSD/aug-cc-pVDZ level of theory ($n_{MO} = 453$). Energies and oscillator strengths are reported in Table II. Timings for e^T 1.0, Dalton 2018, and QChem 5.0 are given in Table III. Thresholds in e^T were set to target an energy convergence of 10^{-6}: the residuals were converged to 10^{-6} for the ground state and 10^{-3} for the excited states (assuming quadratic errors for the energy). In QChem, thresholds for ground and excited states were set to 10^{-6}. We report the total wall time for each calculation. The excited state timing includes the time to converge ground state and excited state equations. The oscillator strength timing also includes the time to solve the multiplier and the left excited state equations. e^T and QChem are equally efficient for the CCSD ground state, while Dalton is considerably slower. For the CCSD excited state calculation, QChem reduced the wall time by a factor of 1.6 compared to e^T and a factor of 5.6 compared to Dalton. For the oscillator strength calculations, QChem reduced the wall time by a factor of 2.7 compared to e^T. The superior performance of QChem for oscillator strengths is primarily due to an efficient starting guess for the left excitation vectors, which are restarted from the right vectors: only 27 transformations are needed to converge all four roots. In e^T 1.0, orbital differences are used as the starting guess for both left and right states, which explains the poorer performance for oscillator strengths.
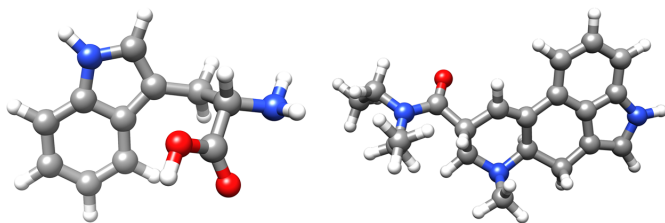
TABLE II. CCSD/aug-cc-pVDZ excitation energies (ω) and oscillator strengths (f_ω) for tryptophan.

	ω (eV)	f_ω
S_1	4.806	0.032
S_2	4.821	0.001
S_3	4.972	0.088
S_4	5.364	0.001

TABLE III. Total calculation times for CCSD/aug-cc-pVDZ ground state (t_{gs}), excitation energy (t_ω), and oscillator strength (t_{f_ω}) calculations for tryptophan. n_{calls}^{gs} is the number of calculations of the residual vector for the ground state. n_{calls}^R and n_{calls}^L are the number of calls to the Jacobian and Jacobian transpose transformations, respectively. The calculations were performed on an Intel Xeon E5-2699 v4 using 44 threads and 1.5 TB shared memory.

	t_{gs} (min)	t_ω (h)	t_{f_ω} (h)	n_{calls}^{gs}	n_{calls}^R	n_{calls}^L
Dalton 2018	1409	84	...	18	88	...
e^T 1.0	201	24	53	16	79	81
QChem 5.0	196	15	20	18	90	27

TABLE IV. The FC-CCSD/aug-cc-pVDZ correlation energy (E_{corr}) and lowest excitation energy (ω) of LSD. A set of decomposition thresholds (τ) for the Cholesky decomposition of the electron repulsion integral matrix were used. Both the ground and excited state equations are converged to within a residual threshold of 10^{-6}. Deviations in the correlation and excitation energies (ΔE_{corr} and $\Delta \omega$) are relative to $\tau = 10^{-8}$.

τ	E_{corr} (E_h)	ΔE_{corr} (E_h)	ω (E_h)	$\Delta \omega$ (E_h)
10^{-2}	$-3.649\,673\,3$	2.3×10^{-2}	$0.165\,734\,3$	7.1×10^{-4}
10^{-3}	$-3.672\,021\,8$	2.3×10^{-4}	$0.165\,037\,0$	7.7×10^{-6}
10^{-4}	$-3.672\,342\,1$	-9.2×10^{-5}	$0.165\,027\,9$	-1.4×10^{-6}
10^{-6}	$-3.672\,254\,2$	-3.6×10^{-6}	$0.165\,029\,4$	1.1×10^{-7}
10^{-8}	$-3.672\,250\,6$	...	$0.165\,029\,3$	...

We have performed FC-CCSD/aug-cc-pVDZ calculations on LSD ($n_{MO} = 777$, $n_{frozen} = 24$). To demonstrate the effect of integral approximation through Cholesky decomposition, we consider a range of decomposition thresholds. The correlation energy and the lowest excitation energy are given in Table IV. Both ground

FIG. 2. Tryptophan (left) and LSD (right).

and excited state residual thresholds are 10^{-6}. With a decomposition threshold of 10^{-2}, the error in the excitation energy ($\Delta \omega$) is less than $10^{-3}E_h$, well within the expected accuracy of FC-CCSD.[74–76]

The CC3 model can be used to obtain highly accurate excitation energies. However, an iterative cost that scales as $\mathcal{O}(n_v^4 n_o^3)$ severely limits system size. To the best of our knowledge, e^T 1.0 includes the fastest available implementation of CC3. A ground and excited state calculation on glycine (chemical formula $C_2H_5NO_2$) using the aug-cc-pVDZ basis set took 33 min with e^T 1.0. Comparable numbers for the new[23] and old[80] CC3 implementations in Dalton 2018[3] were 73 min and 1279 min, respectively.

We have calculated valence and core excitation energies and EOM oscillator strengths for the nucleobase uracil (chemical formula $C_4H_4N_2O_2$, see Fig. 3). The geometry was optimized at the CCSD(T)/aug-cc-pVDZ level using CFOUR.[2] One valence excitation energy was calculated at the FC-CCSD/aug-cc-pVTZ and FC-CC3/aug-cc-pVTZ levels of theory ($n_{MO} = 452$). Two core excited states were calculated for each of the oxygen atoms (O1 and O2, see Fig. 3) at the CCSD and CC3 levels. The aug-cc-pCVTZ basis was used on the oxygen being excited and aug-cc-pVDZ on the remaining atoms ($n_{MO} = 256$). The results are given in Table V. The total timings for the uracil calculations are presented in Table VI. In Table VII, we present the averaged timings from the CVS calculations. They clearly demonstrate the reduced computational cost of the CVS implementation for CC3. The ground state calculation was about four times more expensive per iteration than the right excited state. Without the CVS approximation, the computational cost of the excited states scales as $4n_v^4 n_o^3$ per iteration, while the ground state scales as $2n_v^4 n_o^3$. Using CVS, the excited state scaling is reduced to $4n_v^4 n_o^2$.

In Table VIII, we compare the timings for solving the ground and right excited state equations of glycine with aug-cc-pVDZ for different number of threads. All calculations were run on similar nodes, and all 40 cores on each node were reserved for the jobs to
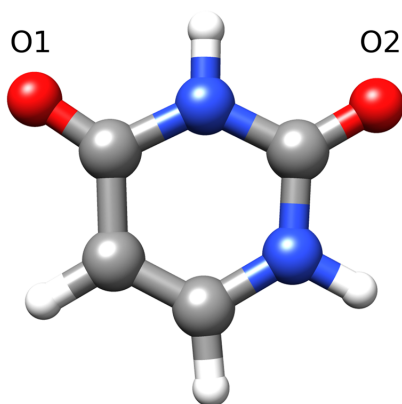
FIG. 3. Uracil with labels on the oxygens.

TABLE V. CC3 valence and core (oxygen edge) excitation energies (ω) and EOM oscillator strengths (f_ω) for uracil. Valence excitations were calculated with the aug-cc-pVTZ basis on all atoms and the frozen core approximation. Core excitations were calculated using the CVS approximation with the aug-cc-pCVTZ basis on the oxygen atom being excited and the aug-cc-pVDZ basis on the remaining atoms.

	CCSD		CC3	
	ω (eV)	f_ω	ω (eV)	f_ω
Valence	5.08	2.24×10^{-8}	4.81	2.23×10^{-6}
Core O1	536.04	3.35×10^{-2}	533.64	1.95×10^{-2}
	539.60	3.23×10^{-4}	535.66	2.24×10^{-4}
Core O2	536.98	3.13×10^{-2}	534.64	1.32×10^{-2}
	539.44	1.47×10^{-4}	535.75	1.34×10^{-4}

TABLE VI. Total wall times for CC3 on uracil. The valence calculation was performed on a node with two Intel Xeon Gold 6138 processors using 40 threads and 320 GB shared memory. The CVS calculations were performed on a node with two Intel Xeon Gold 6138 processors using 40 threads and 150 GB shared memory. n_o and n_v are the number of occupied and virtual orbitals, respectively.

Calculation	Basis set	t (h)	n_o	n_v
Valence excitation	aug-cc-pVTZ	147	21	431
CVS O1	aug-cc-pV(CT)Z	36	29	227
CVS O2	aug-cc-pV(CT)Z	38	29	227

TABLE VII. Average wall time per function call for both CC3 core excitation calculations on uracil. n_{calls} is the total number of routine calls in the two calculations.

Contributions	t (min)	n_{calls}
Ground state amplitudes	14	28
Ground state multipliers	23	30
Right excited states	4	195
Left excited states	7	244

minimize variation. Increasing the number of threads results in significant reductions in time, even for a relatively small system such as glycine with 20 occupied and 140 virtual orbitals. Intermediates are currently stored on disk, resulting in overhead that can be reduced by placing them in memory when possible. In addition to more adaptive memory usage, we are working on improving the coupled cluster algorithms for better parallelization.

B. Cholesky decomposition

We have determined the Cholesky basis for the transmembrane ion channel gramicidin A (chemical formula $C_{198}H_{276}N_{40}O_{34}$, see Fig. 4). The geometry is taken from the supplementary material of Ref. 81. Decomposition times are given in Table IX for the cc-pVDZ and aug-cc-pVDZ basis sets and a range of

TABLE VIII. Time to solve the ground and right excited state equations of glycine for CC3 and CCSD with different numbers of threads in seconds using a development version of e^T 1.1. Factor is the improvement compared to the row above. With perfect parallelization, the factors would be 5, 2, 2, and 2. The calculations were performed on nodes with two Intel Xeon Gold 6138 processors and 150 GB shared memory.

Threads	CC3 GS Time (s)	CC3 GS Factor	CC3 ES Time (s)	CC3 ES Factor	CCSD GS Time (s)	CCSD GS Factor	CCSD ES Time (s)	CCSD ES Factor
1	6048	. . .	15 617	. . .	639	. . .	654	. . .
5	1681	3.60	4 418	3.53	150	4.27	238	2.75
10	923	1.82	2 308	1.91	76	1.96	126	1.89
20	675	1.37	1 482	1.56	41	1.88	72	1.76
40	532	1.27	1 252	1.18	30	1.34	53	1.34

decomposition thresholds. These are compared to the time of one HF iteration. Except when using cc-pVDZ with the tightest threshold, the decomposition time is small or negligible compared to one Fock matrix construction.
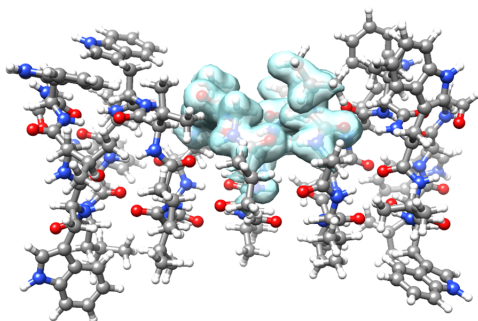
FIG. 4. Gramicidin A. The active MLHF/cc-pVDZ density is shown.

TABLE IX. Cholesky decomposition wall times (t) for gramicidin. τ is the decomposition threshold, and n_J is the number of Cholesky vectors. For reference, we include the time (t_{it}^{HF}) for one full Hartree–Fock iteration. All calculations were performed on an Intel Xeon E5-2699 v4 using 44 threads and 1.5 TB shared memory.

Basis	n_{AO}	τ	n_J	t (min)	t_{it}^{HF} (min)
cc-pVDZ	5188	10^{-2}	11 574	3	35
		10^{-3}	16 368	6	
		10^{-4}	24 652	12	
		10^{-8}	75 446	125	
aug-cc-pVDZ	8740	10^{-2}	12 813	8	1191
		10^{-3}	18 587	27	
		10^{-4}	29 818	61	
		10^{-8}	90 656	645	

C. Hartree–Fock

Systems with several hundred atoms are easily modeled in e^T using Hartree–Fock. In Table X, we present the wall times for calculations on gramicidin A (see Fig. 4) and an amylose chain with 16 glucose units (chemical formula $C_{96}H_{162}O_{81}$, see Fig. 5). The amylose geometry is taken from Ref. 24. We compare the results and timings from e^T 1.0 and QChem 5.0.[11] This comparison is complicated because the accuracy depends on several thresholds apart from the gradient and energy thresholds, e.g., screening thresholds and integral accuracy. Therefore, we list the energies and absolute energy differences along with the timings in Table X. QChem 5.0 outperforms e^T by about a factor of 2. The energies converge to slightly different results in the two programs. In the case of amylose, we find a 2×10^{-7} E_h energy difference using the tightest thresholds ($\tau_{SCF} = 10^{-10}$). Since QChem is a closed source program, we do not know the reason for the deviation. However, we are able to reproduce the e^T results for amylose to all digits using tight thresholds in LSDalton 2018.[3]

D. Multilevel and multiscale methods

To demonstrate the efficacy of multilevel methods for excitation energies, we consider a system of sulfur dioxide with 21 water molecules (see Fig. 6). In Table XI, we present different flavors of multilevel calculations to approximate the two lowest FC-CC2 excitation energies for this system. Three sets of active atoms are defined. The first set contains sulfur dioxide and nine water molecules; these atoms determine the active orbitals of the MLHF calculation. The second set contains sulfur dioxide and five water molecules; these atoms determine the reduced space coupled cluster calculations. The third set contains only sulfur dioxide and determines the CC2 active space in the MLCC2 calculations. The reduced space FC-CC2 calculations are denoted FC-CC2-in-HF and FC-CC2-in-MLHF and similarly for the reduced space FC-MLCC2 calculations. The orbital spaces are partitioned using the Cholesky occupied orbitals and PAOs for the virtual orbitals. In all calculations, the deviation with respect to full FC-CC2 is within the expected error of CC2.[75,76]

In order to assess the performance of the MLHF implementation, we compare full HF and MLHF for gramicidin A and amylose. The active electron densities from the MLHF calculations are shown

TABLE X. Hartree–Fock/cc-pVDZ calculations on amylose and gramicidin. The total wall time is denoted by t, and τ_{SCF} is the Hartree–Fock convergence threshold. We present timings for e^T and QChem along with the computed Hartree–Fock energies (E) and absolute energy differences ($|\Delta E|$) with respect to the calculation with the tightest threshold. Calculations were performed on two Intel Xeon E5-2699 v4 processors using 44 threads and 1.5 TB shared memory.

	e^T				QChem		
	τ_{SCF}	E (E_h)	$|\Delta E|$ (E_h)	t (min)	τ_{SCF}	E (E_h)	$|\Delta E|$ (E_h)	t (min)
Amylose	10^{-3}	$-9792.085\,129\,90$	4×10^{-5}	21	10^{-5}	$-9\,792.085\,350\,39$	2×10^{-4}	9
	10^{-4}	$-9792.085\,178\,33$	5×10^{-6}	31	10^{-6}	$-9\,792.085\,180\,84$	7×10^{-6}	14
	10^{-5}	$-9792.085\,174\,42$	7×10^{-7}	42	10^{-7}	$-9\,792.085\,171\,19$	2×10^{-6}	19
	10^{-6}	$-9792.085\,173\,77$	1×10^{-8}	60	10^{-8}	$-9\,792.085\,173\,23$	4×10^{-7}	26
	10^{-7}	$-9792.085\,173\,76$	$<1 \times 10^{-8}$	78	10^{-9}	$-9\,792.085\,173\,61$	3×10^{-8}	33
	10^{-10}	$-9792.085\,173\,76$	$\ldots$	153	10^{-10}	$-9\,792.085\,173\,58$	$\ldots$	46
Gramicidin	10^{-4}	$-12\,383.458\,832\,54$	4×10^{-6}	130	10^{-6}	$-12\,383.458\,825\,13$	1×10^{-5}	50
	10^{-5}	$-12\,383.458\,836\,34$	7×10^{-8}	198	10^{-7}	$-12\,383.458\,827\,10$	1×10^{-5}	77
	10^{-6}	$-12\,383.458\,836\,27$	$\ldots$	280	10^{-8}	$-12\,383.458\,836\,77$	$\ldots$	111
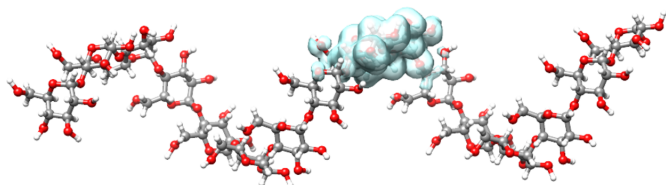
FIG. 5. Amylose chain of 16 glucose units. The active MLHF/cc-pVDZ density is shown.
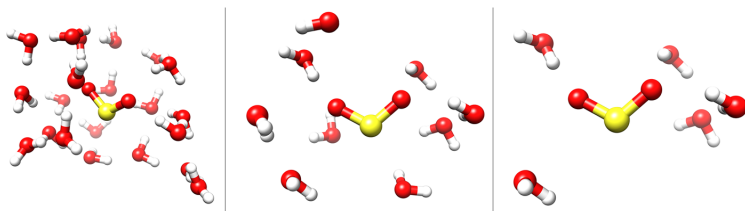
FIG. 6. SO_2 and water. (Left) SO_2 and 21 water molecules. (Middle) SO_2 and nine water molecules; these are the HF active atoms in the MLHF calculations. (Right) SO_2 and five water molecules; these are the CC active atoms. In the MLCC2 calculations, only SO_2 is treated at the CC2 level of theory.

TABLE XI. The two lowest excitation energies (ω_1 and ω_2) of SO_2 with 21 water molecules, calculated with full and reduced space FC-CC2 and FC-MLCC2 using HF and MLHF reference wave functions. The deviation from full FC-CC2 ($\Delta\omega_i = \omega_i - \omega_i^{FC-CC2}$) is given. We also list the number of occupied (n_o) and virtual (n_v) orbitals treated at the different levels of theory. There are a total of 121 occupied orbitals and 813 virtual orbitals in the system.

	HF		CCS		CC2		ω_1	$\Delta\omega_1$	ω_2	$\Delta\omega_1$
Calculation	n_o	n_v	n_o	n_v	n_o	n_v	(eV)	(eV)	(eV)	(eV)
FC-CC2	121	813	$\ldots$	$\ldots$	93	813	3.11	$\ldots$	3.39	$\ldots$
FC-CC2-in-HF	121	813	$\ldots$	$\ldots$	40	266	3.14	0.03	3.43	0.04
FC-CC2-in-MLHF	75	426	$\ldots$	$\ldots$	40	266	3.16	0.05	3.44	0.05
FC-MLCC2	121	813	93	813	14	67	3.18	0.07	3.45	0.06
FC-MLCC2-in-HF	121	813	40	266	14	66	3.18	0.07	3.45	0.06
FC-MLCC2-in-MLHF	75	426	40	266	15	66	3.20	0.09	3.47	0.08

TABLE XII. Multilevel Hartree–Fock wall times for amylose and gramicidin. t_{fit} is the wall time to construct the Fock matrix. For the calculations with (aug)-cc-pVDZ, aug-cc-pVDZ is used on the active atoms and cc-pVDZ for the rest. The total number of AOs and the active MOs is labeled n_{AO} and n_{MO}^{active}, respectively. Thresholds for Coulomb and exchange are set to 10^{-12} and 10^{-10}, respectively, and the integral cutoff is set to 10^{-12}. Calculations were performed on two Intel Xeon E5-2699 v4 processors using 44 threads and 1.5 TB shared memory.

		HF		MLHF	
	Basis	n_{AO}	t_{it} (min)	n_{MO}^{active}	t_{it} (min)
Amylose	cc-pVDZ	3288	8	335	1
	(aug)-cc-pVDZ	3480	11	552	4
Gramicidin	cc-pVDZ	5188	35	546	11
	(aug)-cc-pVDZ	5506	69	942	50

in Figs. 4 and 5. The plots were generated using UCSF Chimera.[82] Cholesky orbitals were used to partition the occupied space, and PAOs were used for the virtual space. We present the timings in Table XII. For amylose, the iteration times are reduced significantly with MLHF: by a factor of 8 when cc-pVDZ is used on all atoms

and a factor of 3 when aug-cc-pVDZ is used on the active atoms. In contrast, only a factor of 3 was reported by Sæther *et al.*[24] in the cc-pVDZ case. The iteration time is also reduced by a factor of 8 for amylose/cc-pVDZ ($t_{iteration}$ = 1 m, n_{MO}^{active} = 318) when using Cholesky virtuals (as in Ref. 24) instead of PAOs. The savings for amylose reflect the small active region as well as the linear structure of the chain. Savings are less significant for the gramicidin system, where the MLHF iteration time is a third of the HF iteration time for cc-pVDZ, but only about two thirds when the active atoms are described using aug-cc-pVDZ. The smaller savings reflect the relatively large active region and the more compact shape of the gramicidin system.

For systems in solution, electronic spectra can be calculated using QM/MM or QM/PCM. Paranitroaniline (PNA) has an experimental vacuum-to-water solvatochromism of about 1 eV.[87] For QM/PCM, we use two different atomic radii, UFF[85] (QM/PCM[c]) and Bondi[86] (QM/PCM[d]), and the dielectric permittivity of water was set to ε = 78.39. For QM/MM, 64 snapshots were extracted from a classical molecular dynamics simulation[88] [see Fig. 7(a) for an example structure]. The UV/vis spectra were then computed by treating PNA at the CC2/aug-cc-pVDZ level and modeling the water using an FQ force field. Here, we present results using two different FQ parameterizations: QM/FQ[a] from Ref. 83 and QM/FQ[b] from
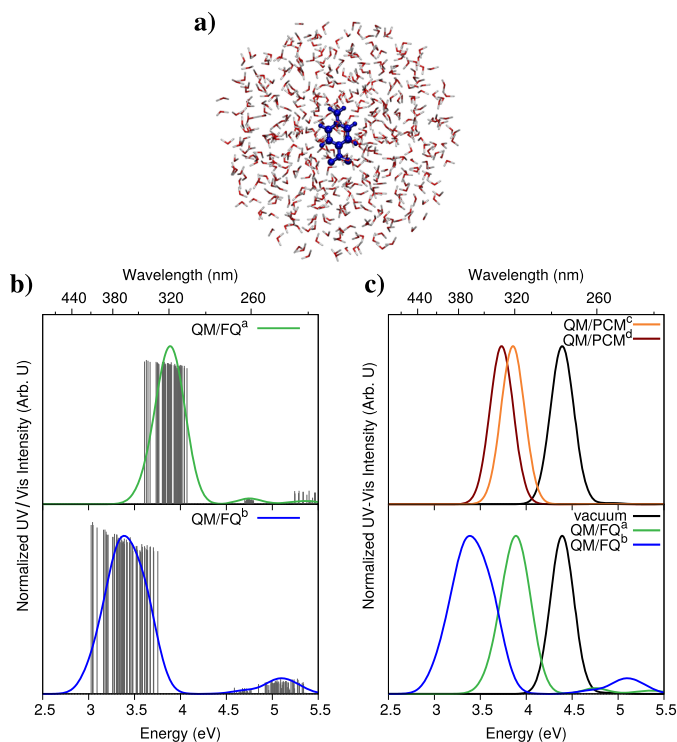
FIG. 7. (a) Schematic representation of a random snapshot of PNA in aqueous solution. (b) and (c) UV/vis spectra of PNA calculated at the CC2/aug-cc-pVDZ level of theory with an aqueous solution described at the PCM or FQ level of theory. (b) QM/FQ raw data (sticks) together with their Gaussian convolution (FWHM = 0.3 eV). (c) QM/PCM (top) and QM/FQ (bottom) spectra in aqueous solution. A gas phase CC2/aug-cc-pVDZ reference spectrum is also reported (black). For QM/FQ[a], the FQ parameterization is from Ref. 83, and for QM/FQ[b], the parameterization is from Ref. 84. In QM/FQ[c], the PCM cavity is constructed using the UFF radii,[85] and in QM/FQ[d], it is constructed using the Bondi radii.[86]

TABLE XIII. The first vertical excitation energy of PNA in vacuum (ω_v) and in aqueous solution (ω_s), as well as water-to-vacuum solvatochromic shifts ($\Delta\omega$). PNA is treated at the CC2/aug-cc-pVDZ level of theory, and the solution is described with PCM or FQ. 68% confidence intervals for excitation energies are also reported for QM/FQ, calculated as $\sigma/\sqrt{N}$, where σ is the standard deviation and N is the number of the snapshots used to obtain the average property. Experimental data are taken from Ref. 87.

	ω_v (eV)	ω_s (eV)	$\Delta\omega$ (eV)
CC2	4.38	...	...
CC2/FQ[a]	...	3.88±0.01	0.50±0.01
CC2/FQ[b]	...	3.38±0.01	1.00±0.01
CC2/PCM[c]	...	3.86	0.52
CC2/PCM[d]	...	3.76	0.62
Expt.[e]	4.25	3.26	0.99

[a]FQ parameterization taken from Ref. 83.
[b]FQ parameterization taken from Ref. 84.
[c]PCM cavity constructed by exploiting UFF radii.[85]
[d]PCM cavity constructed by exploiting the Bondi radii.[86]
[e]Reference 87.

Ref. 84 (see the supplementary material for additional computational details).

The spectra calculated using QM/FQ are presented in Fig. 7(b). The results for individual snapshots are presented as sticks together with their Gaussian convolution. As can be seen from Fig. 7, QM/FQ[b] results in a greater spread in the excitation energies. This

is probably due to the larger molecular dipole moments of the water molecules in this parameterization.[88,89]

In Fig. 7(c), the convoluted spectra calculated using QM/PCM[c] and QM/PCM[d] (top), and QM/FQ[a] and QM/FQ[b] (bottom), are presented with their vacuum counterparts. The excitation energies are also given in Table XIII together with the experimental data from Ref. 87. For QM/FQ, we also report 68% confidence intervals for the calculated excitation energies. QM/FQ[b] reproduces the experimental solvatochromism, while the other approaches give errors of 40%–50%.

E. Modeling spectroscopies

The spectroscopic properties can also be modeled with the Lanczos method or with real-time propagation of the coupled cluster wave function. In Fig. 8, we show CCSD/aug-cc-pCVDZ UV/vis absorption spectra of H_2O,[90] calculated using the Davidson (top) and asymmetric Lanczos (bottom) algorithms. Note that we have artificially extended the spectra beyond the ionization potential (12.3 eV IP-CCSD/aug-cc-pCVDZ) to illustrate convergence behavior. With the Lanczos algorithm, the low energy part of the spectrum converges with a smaller reduced space than the high energy part.[47]

We have also generated oxygen edge x-ray absorption spectra using the Davidson and Lanczos algorithms with CVS projection (see Fig. 9. We see the same overall behavior as in Fig. 8.

Absorption spectra can also be obtained from real-time propagation of the coupled cluster wave function (see Fig. 10 for UV/vis and oxygen edge x-ray absorption spectra; see the
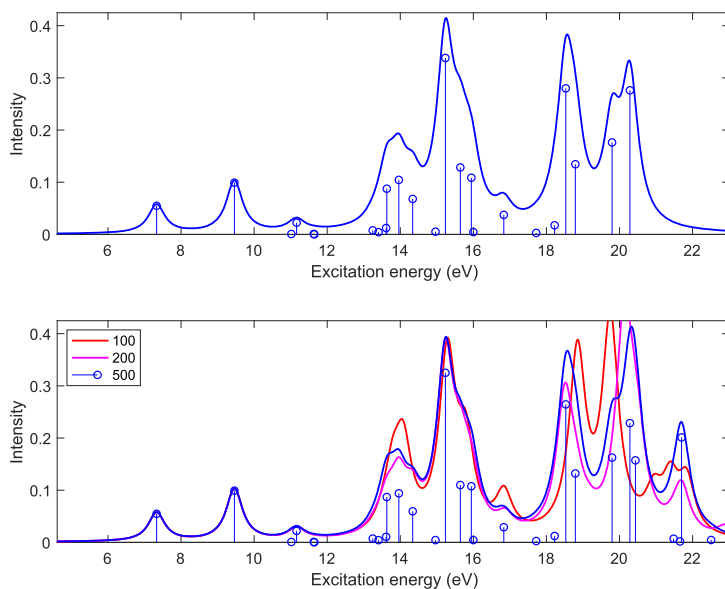
FIG. 8. Water CCSD/aug-cc-pCVDZ UV/vis absorption spectrum. Lorentzian broadening (0.02 E_h FWHM) has been applied to the stick spectra. The top plot shows the spectrum obtained using the Davidson. The spectrum in the bottom plot is from Lanczos calculations with chain lengths 100 (red), 200 (magenta), and 500 (blue).
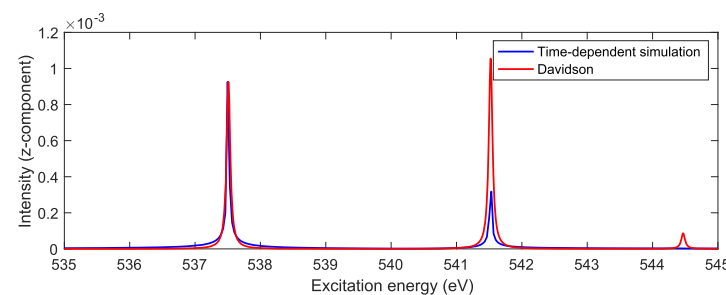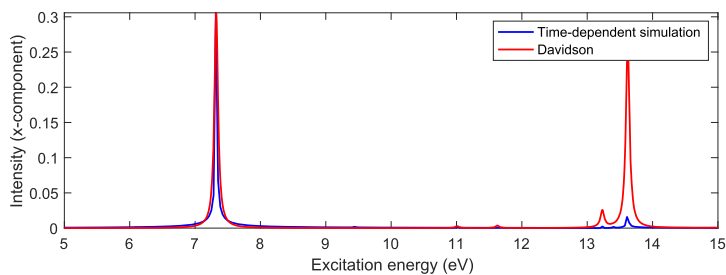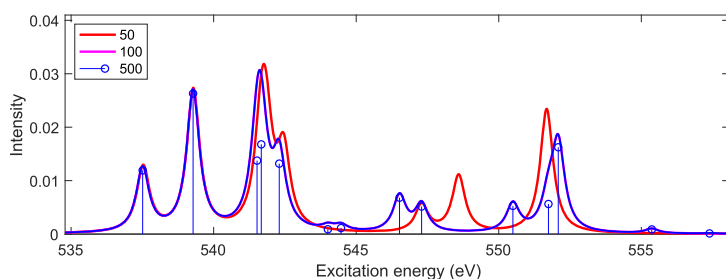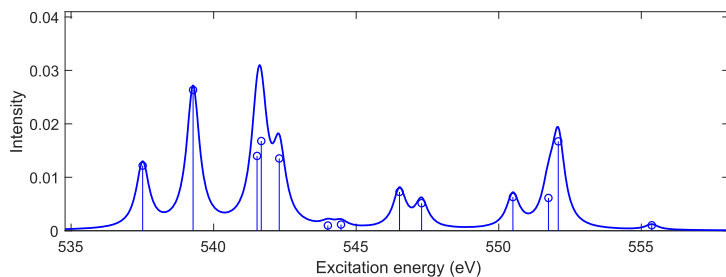
FIG. 9. Water CCSD/CVS/aug-cc-pCVDZ x-ray absorption spectrum. Lorentzian broadening (0.02 E_h FWHM) has been applied to the stick spectra. The top plot shows the spectrum obtained using the Davidson. The spectrum in the bottom plot is from Lanczos calculations with chain lengths 50 (red), 100 (magenta), and 500 (blue).

FIG. 10. Water UV and x-ray CCSD absorption spectra obtained using Davidson (CVS/aug-cc-pCVDZ for x ray and aug-cc-pVDZ for UV) and real-time propagation (aug-cc-pCVDZ for x ray and aug-cc-pVDZ for UV). The top and bottom plots show the simulated UV and x-ray spectra, respectively. The Davidson spectra were produced by applying Lorentzian broadening to the stick spectra (0.0025 E_h FWHM). Intensities from the time-dependent simulation have been scaled so that the intensity of the first peak matches the EOM oscillator strength.

supplementary material for computational details). The first peak in both plots has been scaled to match the intensity obtained using Davidson. The position of the peaks are the same with both approaches, but the intensities differ because we specified pulses with frequency distributions centered on the first excitation energy.

IV. CONCLUDING REMARKS

e^T 1.0 is an optimized open source electronic structure program. Several features are worth emphasizing. To the best of our knowledge, our CC3 implementation is the fastest for calculating

ground and excited state energies and EOM oscillator strengths. The low memory CC2 code has memory and disk requirements of order $\mathcal{O}(n_{\text{MO}}^2)$ and $\mathcal{O}(n_{\text{MO}}^3)$, respectively, allowing us to treat systems with thousands of basis functions. At the core of our program is the Cholesky decomposition of the electron repulsion integral matrix; our implementation is faster and less storage intensive than that of any other program. Exciting new developments are also part of e^T. It features the only spin adapted closed shell implementation of time-dependent coupled cluster theory. Furthermore, the MLHF and MLCC2 methods extend the treatable system size without sacrificing accuracy for intensive properties such as excitation energies.

The e^T source code is written in modern object oriented Fortran, making it easy to expand and contribute to the program. It is freely available on GitLab,[91] and the manual can be found at www.etprogram.org. We will continue to expand the capabilities of e^T, focusing on molecular properties and multilevel methods. We believe that the program will be useful for the quantum chemistry community, both as a development platform and for production calculations.

SUPPLEMENTARY MATERIAL

See the supplementary material for details regarding QM/MM calculations as well as specifications for the time-dependent CCSD propagation calculations.

AUTHOR'S CONTRIBUTIONS

S.D.F, E.F.K., and R.H.M contributed equally to this work.

ACKNOWLEDGMENTS

DATA AVAILABILITY

The e^T source code is publicly available on Gitlab (see Ref. 91). The geometries used in Sec. III are available at https://doi.org/10.5281/zenodo.3666109 (Ref. 73). Input files are available from the corresponding author upon request.

REFERENCES

[1] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic-Structure Theory* (John Wiley & Sons, 2014).

[2] J. Stanton, J. Gauss, M. Harding, P. Szalay, A. Auer, R. Bartlett, U. Benedikt, C. Berger, D. Bernholdt, Y. Bomble, *et al.*, "CFOUR, coupled-cluster techniques for computational chemistry, a quantum-chemical program package," current version available at http://www.cfour.de, 2010.

[3] K. Aidas, C. Angeli, K. L. Bak, V. Bakken, R. Bast, L. Boman, O. Christiansen, R. Cimiraglia, S. Coriani, P. Dahle, E. K. Dalskov, U. Ekström, T. Enevoldsen, J. J. Eriksen, P. Ettenhuber, B. Fernández, L. Ferrighi, H. Fliegl, L. Frediani, K. Hald, A. Halkier, C. Hättig, H. Heiberg, T. Helgaker, A. C. Hennum, H. Hettema, E. Hjertenaes, S. Høst, I.-M. Høyvik, M. F. Iozzi, B. Jansík, H. J. A. Jensen, D. Jonsson, P. Jørgensen, J. Kauczor, S. Kirpekar, T. Kjaergaard, W. Klopper, S. Knecht, R. Kobayashi, H. Koch, J. Kongsted, A. Krapp, K. Kristensen, A. Ligabue, O. B. Lutnaes, J. I. Melo, K. V. Mikkelsen, R. H. Myhre, C. Neiss, C. B. Nielsen, P. Norman, J. Olsen, J. M. H. Olsen, A. Osted, M. J. Packer, F. Pawlowski, T. B. Pedersen, P. F. Provasi, S. Reine, Z. Rinkevicius, T. A. Ruden, K. Ruud, V. V. Rybkin, P. Sałek, C. C. M. Samson, A. S. de Merás, T. Saue, S. P. A. Sauer, B. Schimmelpfennig, K. Sneskov, A. H. Steindal, K. O. Sylvester-Hvid, P. R. Taylor, A. M. Teale, E. I. Tellgren, D. P. Tew, A. J. Thorvaldsen, L. Thøgersen, O. Vahtras, M. A. Watson, D. J. Wilson, M. Ziolkowski, and H. Ågren, "The Dalton quantum chemistry program system," Wiley Interdiscip. Rev.: Comput. Mol. Sci. **4**, 269–284 (2014).

[4] M. S. Gordon and M. W. Schmidt, "Advances in electronic structure theory: GAMESS a decade later," in *Theory and Applications of Computational Chemistry*, edited by C. E. Dykstra, G. Frenking, K. S. Kim, and G. E. Scuseria (Elsevier, Amsterdam, 2005), Chap. 41, pp. 1167–1189.

[5] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian 16, Revision C.01, Gaussian, Inc., Wallingford, CT, 2016.

[6] F. Aquilante, J. Autschbach, R. K. Carlson, L. F. Chibotaru, M. G. Delcey, L. De Vico, I. Fdez. Galván, N. Ferré, L. M. Frutos, L. Gagliardi, M. Garavelli, A. Giussani, C. E. Hoyer, G. Li Manni, H. Lischka, D. Ma, P. Å. Malmqvist, T. Müller, A. Nenov, M. Olivucci, T. B. Pedersen, D. Peng, F. Plasser, B. Pritchard, M. Reiher, I. Rivalta, I. Schapiro, J. Segarra-Martí, M. Stenrup, D. G. Truhlar, L. Ungur, A. Valentini, S. Vancoillie, V. Veryazov, V. P. Vysotskiy, O. Weingart, F. Zapata, and R. Lindh, "MOLCAS 8: New capabilities for multiconfigurational quantum chemical calculations across the periodic table," J. Comput. Chem. **37**, 506–541 (2016).

[7] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, and M. Schütz, "MOLPRO: A general-purpose quantum chemistry program package," Wiley Interdiscip. Rev.: Comput. Mol. Sci. **2**, 242–253 (2012).

[8] M. Valiev, E. J. Bylaska, N. Govind, K. Kowalski, T. P. Straatsma, H. J. J. Van Dam, D. Wang, J. Nieplocha, E. Apra, T. L. Windus, and W. A. de Jong, "NWChem: A comprehensive and scalable open-source solution for large scale molecular simulations," Comput. Phys. Commun. **181**, 1477–1489 (2010).

[9] F. Neese, "The ORCA program system," Wiley Interdiscip. Rev.: Comput. Mol. Sci. **2**, 73–78 (2012).

[10] R. M. Parrish, L. A. Burns, D. G. A. Smith, A. C. Simmonett, A. E. DePrince, E. G. Hohenstein, U. Bozkaya, A. Y. Sokolov, R. Di Remigio, R. M. Richard, J. F. Gonthier, A. M. James, H. R. McAlexander, A. Kumar, M. Saitow, X. Wang, B. P. Pritchard, P. Verma, H. F. Schaefer, K. Patkowski, R. A. King, E. F. Valeev, F. A. Evangelista, J. M. Turney, T. D. Crawford, and C. D. Sherrill, "PSI4 1.1: An open-source electronic structure program emphasizing automation, advanced libraries, and interoperability," J. Chem. Theor. Comput. **13**, 3185–3197 (2017).

[11]Y. Shao, Z. Gan, E. Epifanovsky, A. T. B. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P. R. Horn, L. D. Jacobson, I. Kaliman, R. Z. Khaliullin, T. Kuś, A. Landau, J. Liu, E. I. Proynov, Y. M. Rhee, R. M. Richard, M. A. Rohrdanz, R. P. Steele, E. J. Sundstrom, H. L. Woodcock III, P. M. Zimmerman, D. Zuev, B. Albrecht, E. Alguire, B. Austin, G. J. O. Beran, Y. A. Bernard, E. Berquist, K. Brandhorst, K. B. Bravaya, S. T. Brown, D. Casanova, C.-M. Chang, Y. Chen, S. H. Chien, K. D. Closser, D. L. Crittenden, M. Diedenhofen, R. A. DiStasio, Jr., H. Do, A. D. Dutoi, R. G. Edgar, S. Fatehi, L. Fusti-Molnar, A. Ghysels, A. Golubeva-Zadorozhnaya, J. Gomes, M. W. D. Hanson-Heine, P. H. P. Harbach, A. W. Hauser, E. G. Hohenstein, Z. C. Holden, T.-C. Jagau, H. Ji, B. Kaduk, K. Khistyaev, J. Kim, J. Kim, R. A. King, P. Klunzinger, D. Kosenkov, T. Kowalczyk, C. M. Krauter, K. U. Lao, A. D. Laurent, K. V. Lawler, S. V. Levchenko, C. Y. Lin, F. Liu, E. Livshits, R. C. Lochan, A. Luenser, P. Manohar, S. F. Manzer, S.-P. Mao, N. Mardirossian, A. V. Marenich, S. A. Maurer, N. J. Mayhall, E. Neuscamman, C. M. Oana, R. Olivares-Amaya, D. P. O'Neill, J. A. Parkhill, T. M. Perrine, R. Peverati, A. Prociuk, D. R. Rehn, E. Rosta, N. J. Russ, S. M. Sharada, S. Sharma, D. W. Small, A. Sodt, T. Stein, D. Stück, Y.-C. Su, A. J. W. Thom, T. Tsuchimochi, V. Vanovschi, L. Vogt, O. Vydrov, T. Wang, M. A. Watson, J. Wenzel, A. White, C. F. Williams, J. Yang, S. Yeganeh, S. R. Yost, Z.-Q. You, I. Y. Zhang, X. Zhang, Y. Zhao, B. R. Brooks, G. K. L. Chan, D. M. Chipman, C. J. Cramer, W. A. Goddard III, M. S. Gordon, W. J. Hehre, A. Klamt, H. F. Schaefer III, M. W. Schmidt, C. D. Sherrill, D. G. Truhlar, A. Warshel, X. Xu, A. Aspuru-Guzik, R. Baer, A. T. Bell, N. A. Besley, J.-D. Chai, A. Dreuw, B. D. Dunietz, T. R. Furlani, S. R. Gwaltney, C.-P. Hsu, Y. Jung, J. Kong, D. S. Lambrecht, W. Liang, C. Ochsenfeld, V. A. Rassolov, L. V. Slipchenko, J. E. Subotnik, T. V. Voorhis, J. M. Herbert, A. I. Krylov, P. M. W. Gill, and M. Head-Gordon, "Advances in molecular quantum chemistry contained in the Q-Chem 4 program package," Mol. Phys. **113**, 184–215 (2015).

[12]F. Furche, R. Ahlrichs, C. Hättig, W. Klopper, M. Sierka, and F. Weigend, "Turbomole," Wiley Interdiscip. Rev.: Comput. Mol. Sci. **4**, 91–100 (2014).

[13]T. Helgaker, S. Coriani, P. Jørgensen, K. Kristensen, J. Olsen, and K. Ruud, "Recent advances in wave function-based methods of molecular-property calculations," Chem. Rev. **112**, 543–631 (2012).

[14]H. Koch, H. J. A. Jensen, P. Jørgensen, T. Helgaker, G. E. Scuseria, and H. F. Schaefer, "Coupled cluster energy derivatives. Analytic hessian for the closed-shell coupled cluster singles and doubles wave function: Theory and applications," J. Chem. Phys. **92**, 4924–4940 (1990).

[15]J. F. Stanton, "Many-body methods for excited state potential energy surfaces. I. General theory of energy gradients for the equation-of-motion coupled-cluster method," J. Chem. Phys. **99**, 8840–8847 (1993).

[16]J. F. Stanton and R. J. Bartlett, "The equation of motion coupled-cluster method. A systematic biorthogonal approach to molecular excitation energies, transition probabilities, and excited state properties," J. Chem. Phys. **98**, 7029–7039 (1993).

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

[18]N. H. F. Beebe and J. Linderberg, "Simplifications in the generation and transformation of two-electron integrals in molecular calculations," Int. J. Quantum Chem. **12**, 683–705 (1977).

[19]H. Koch, A. Sánchez de Merás, and T. B. Pedersen, "Reduced scaling in electronic structure calculations using Cholesky decompositions," J. Chem. Phys. **118**, 9481–9484 (2003).

[20]F. Coester and H. Kümmel, "Short-range correlations in nuclear wave functions," Nucl. Phys. **17**, 477–485 (1960).

[21]S. D. Folkestad, E. F. Kjønstad, and H. Koch, "An efficient algorithm for Cholesky decomposition of electron repulsion integrals," J. Chem. Phys. **150**, 194112 (2019).

[22]H. Koch, O. Christiansen, P. Jørgensen, A. M. Sanchez de Merás, and T. Helgaker, "The CC3 model: An iterative coupled cluster approach including connected triples," J. Chem. Phys. **106**, 1808–1818 (1997).

[23]R. H. Myhre and H. Koch, "The multilevel CC3 coupled cluster model," J. Chem. Phys. **145**, 044111 (2016).

[24]S. Sæther, T. Kjærgaard, H. Koch, and I.-M. Høyvik, "Density-based multilevel Hartree–Fock model," J. Chem. Theor. Comput. **13**, 5282–5290 (2017).

[25]R. H. Myhre, A. M. J. Sánchez de Merás, and H. Koch, "Multi-level coupled cluster theory," J. Chem. Phys. **141**, 224105 (2014).

[26]S. D. Folkestad and H. Koch, "The multilevel CC2 and CCSD methods with correlated natural transition orbitals," J. Chem. Theor. Comput. **16**, 179–189 (2019).

[27]A. Warshel and M. Karplus, "Calculation of ground and excited state potential surfaces of conjugated molecules. I. Formulation and parametrization," J. Am. Chem. Soc. **94**, 5612–5625 (1972).

[28]M. Levitt and A. Warshel, "Computer simulation of protein folding," Nature **253**, 694–698 (1975).

[29]J. Tomasi, B. Mennucci, and R. Cammi, "Quantum mechanical continuum solvation models," Chem. Rev. **105**, 2999–3094 (2005).

[30]B. Mennucci, "Polarizable continuum model," Wiley Interdiscip. Rev.: Comput. Mol. Sci. **2**, 386–404 (2012).

[31]E. Valeev, "Libint: A library for the evaluation of molecular integrals of many-body operators over Gaussian functions, Version 2.7.0" (2017), available at http://libint.valeyev.net/.

[32]R. Di Remigio, A. H. Steindal, K. Mozgawa, V. Weijo, H. Cao, and L. Frediani, "PCMSolver: An open-source library for solvation modeling," Int. J. Quantum Chem. **119**, e25685 (2019).

[33]R. Bast (2018). "Runtest," Zenodo. https://doi.org/10.5281/zenodo.1434751

[34]R. Bast, R. Di Remigio, and J. Juselius (2020). "Autocmake," Zenodo. https://doi.org/10.5281/zenodo.3634941

[35]See https://www.openmp.org/spec-html/5.0/openmp.html for technical specifications for OMP.

[36]O. Christiansen, H. Koch, and P. Jørgensen, "The second-order approximate coupled cluster singles and doubles model CC2," Chem. Phys. Lett. **243**, 409–418 (1995).

[37]G. D. Purvis and R. J. Bartlett, "A full coupled-cluster singles and doubles model: The inclusion of disconnected triples," J. Chem. Phys. **76**, 1910–1918 (1982).

[38]K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, "A fifth-order perturbation comparison of electron correlation theories," Chem. Phys. Lett. **157**, 479–483 (1989).

[39]E. R. Davidson, "The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices," J. Comput. Phys. **17**, 87–94 (1975).

[40]P. Pulay, "Improved SCF convergence acceleration," J. Comput. Chem. **3**, 556–560 (1982).

[41]P. Pulay, "Convergence acceleration of iterative sequences. the case of SCF iteration," Chem. Phys. Lett. **73**, 393–398 (1980).

[42]G. E. Scuseria, T. J. Lee, and H. F. Schaefer, "Accelerating the convergence of the coupled-cluster approach: The use of the DIIS method," Chem. Phys. Lett. **130**, 236–239 (1986).

[43]C. Hättig and F. Weigend, "CC2 excitation energy calculations on large molecules using the resolution of the identity approximation," J. Chem. Phys. **113**, 5154–5161 (2000).

[44]M. Ziółkowski, V. Weijo, P. Jørgensen, and J. Olsen, "An efficient algorithm for solving nonlinear equations with a minimal number of trial vectors: Applications to atomic-orbital based coupled-cluster theory," J. Chem. Phys. **128**, 204105 (2008).

[45]P. Ettenhuber and P. Jørgensen, "Discarding information from previous iterations in an optimal way to solve the coupled cluster amplitude equations," J. Chem. Theor. Comput. **11**, 1518–1524 (2015).

[46]C. Lanczos, *An Iteration Method for the Solution of the Eigenvalue Problem of Linear Differential and Integral Operators* (United States Government Press Office, Los Angeles, CA, 1950).

[47]S. Coriani, T. Fransson, O. Christiansen, and P. Norman, "Asymmetric-Lanczos-chain-driven implementation of electronic resonance convergent coupled-cluster linear response theory," J. Chem. Theor. Comput. **8**, 1616–1628 (2012).

[48]H. Koch and P. Jørgensen, "Coupled cluster response functions," J. Chem. Phys. **93**, 3333–3344 (1990).

[49]T. B. Pedersen and S. Kvaal, "Symplectic integration and physical interpretation of time-dependent coupled-cluster theory," J. Chem. Phys. **150**, 144106 (2019).

[50]P. N. Swarztrauber, "FFT algorithms for vector computers," Parallel Comput. **1**, 45–63 (1984).

[51] F. Aquilante, L. Boman, J. Boström, H. Koch, R. Lindh, A. S. de Merás, and T. B. Pedersen, "Cholesky decomposition techniques in electronic structure theory," in *Linear-Scaling Techniques in Computational Chemistry and Physics* (Springer, 2011), pp. 301–343.

[52] J. H. Van Lenthe, R. Zwaans, H. J. J. Van Dam, and M. F. Guest, "Starting SCF calculations by superposition of atomic densities," J. Comput. Chem. **27**, 926–932 (2006).

[53] A. M. J. Sánchez de Merás, H. Koch, I. G. Cuesta, and L. Boman, "Cholesky decomposition-based definition of atomic subsystems in electronic structure calculations," J. Chem. Phys. **132**, 204105 (2010).

[54] P. Pulay, "Localizability of dynamic electron correlation," Chem. Phys. Lett. **100**, 151–154 (1983).

[55] S. Saebø and P. Pulay, "Local treatment of electron correlation," Ann. Rev. Phys. Chem. **44**, 213–236 (1993).

[56] P.-O. Löwdin, "On the nonorthogonality problem," in *Advances in Quantum Chemistry* (Elsevier, 1970), Vol. 5, pp. 185–199.

[57] I.-M. Høyvik, "Convergence acceleration for the multilevel Hartree–Fock model," Mol. Phys. **118**, 1626929 (2020).

[58] R. H. Myhre, A. M. J. Sánches de Merás, and H. Koch, "The extended CC2 model ECC2," Mol. Phys. **111**, 1109–1118 (2013).

[59] I.-M. Høyvik, R. H. Myhre, and H. Koch, "Correlated natural transition orbitals for core excitation energies in multilevel coupled cluster models," J. Chem. Phys. **146**, 144109 (2017).

[60] P. Baudin and K. Kristensen, "Correlated natural transition orbital framework for low-scaling excitation energy calculations (CorNFLEx)," J. Chem. Phys. **146**, 214114 (2017).

[61] H. M. Senn and W. Thiel, "QM/MM methods for biomolecular systems," Angew. Chem., Int. Ed. **48**, 1198–1229 (2009).

[62] C. Cappelli, "Integrated QM/polarizable MM/continuum approaches to model chiroptical properties of strongly interacting solute-solvent systems," Int. J. Quantum Chem. **116**, 1532–1542 (2016).

[63] F. Lipparini, C. Cappelli, and V. Barone, "Linear response theory and electronic transition energies for a fully polarizable QM/classical Hamiltonian," J. Chem. Theory Comput. **8**, 4153–4165 (2012).

[64] R. Di Remigio, T. Giovannini, M. Ambrosetti, C. Cappelli, and L. Frediani, "Fully polarizable QM/fluctuating charge approach to two-photon absorption of aqueous solutions," J. Chem. Theory Comput. **15**, 4056–4068 (2019).

[65] M. Caricato, "Coupled cluster theory with the polarizable continuum model of solvation," Int. J. Quantum Chem. **119**, e25710 (2019).

[66] R. Cammi, "Quantum cluster theory for the polarizable continuum model. I. The CCSD level with analytical first and second derivatives," J. Chem. Phys. **131**, 164104 (2009).

[67] M. Caricato, "Absorption and emission spectra of solvated molecules with the EOM–CCSD–PCM method," J. Chem. Theory Comput. **8**, 4494–4502 (2012).

[68] M. Caricato, F. Lipparini, G. Scalmani, C. Cappelli, and V. Barone, "Vertical electronic excitations in solution with the EOM–CCSD method combined with a polarizable explicit/implicit solvent model," J. Chem. Theory Comput. **9**, 3035–3042 (2013).

[69] S. Ren, F. Lipparini, B. Mennucci, and M. Caricato, "Coupled cluster theory with induced dipole polarizable embedding for ground and excited states," J. Chem. Theory Comput. **15**, 4485–4496 (2019).

[70] L. S. Cederbaum, W. Domcke, and J. Schirmer, "Many-body theory of core holes," Phys. Rev. A **22**, 206–222 (1980).

[71] S. Coriani and H. Koch, "Communication: X-ray absorption spectra and core-ionization potentials within a core-valence separated coupled cluster framework," J. Chem. Phys. **143**, 181103 (2015).

[72] S. Coriani and H. Koch, "Erratum: "Communication: X-ray absorption spectra and core-ionization potentials within a core-valence separated coupled cluster framework" [J. Chem. Phys. 143, 181103 (2015)]," J. Chem. Phys. **145**, 149901 (2016).

[73] J. H. Andersen, A. Balbi, S. Coriani, S. D. Folkestad, T. Giovannini, L. Goletto, T. S. Haugland, A. Hutcheson, I.-M. Høyvik, E. F. Kjønstad, T. Moitra, R. H. Myhre, A. C. Paul, M. Scavino, A. S. Skeidsvoll, Å. H. Tveten, and H. Koch (2020). "Geometries e^T 1.0 paper," Zenodo. https://doi.org/10.5281/zenodo.3666109

[74] O. Christiansen, H. Koch, P. Jørgensen, and J. Olsen, "Excitation energies of H_2O, N_2 and C_2 in full configuration interaction and coupled cluster theory," Chem. Phys. Lett. **256**, 185–194 (1996).

[75] D. Kánnár and P. G. Szalay, "Benchmarking coupled cluster methods on valence singlet excited states," J. Chem. Theor. Comput. **10**, 3757–3765 (2014).

[76] D. Kánnár, A. Tajti, and P. G. Szalay, "Accuracy of coupled cluster excitation energies in diffuse basis sets," J. Chem. Theor. Comput. **13**, 202–209 (2016).

[77] See https://pubchem.ncbi.nlm.nih.gov/compound/Rifampicin for Rifampicin, 2020.

[78] See https://pubchem.ncbi.nlm.nih.gov/compound/Tryptophan for Tryptophan, 2020.

[79] See https://pubchem.ncbi.nlm.nih.gov/compound/Lysergide for Lysergide, 2020.

[80] K. Hald, P. Jørgensen, O. Christiansen, and H. Koch, "Implementation of electronic ground states and singlet and triplet excitation energies in coupled cluster theory with approximate triples corrections," J. Chem. Phys. **116**, 5963–5970 (2002).

[81] S. Li, W. Li, and T. Fang, "An efficient fragment-based approach for predicting the ground-state energies and structures of large molecules," J. Am. Chem. Soc. **127**, 7215–7226 (2005).

[82] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin, "UCSF Chimera–A visualization system for exploratory research and analysis," J. Comput. Chem. **25**, 1605–1612 (2004).

[83] S. W. Rick, S. J. Stuart, and B. J. Berne, "Dynamical fluctuating charge force fields: Application to liquid water," J. Chem. Phys. **101**, 6141–6156 (1994).

[84] T. Giovannini, P. Lafiosca, B. Chandramouli, V. Barone, and C. Cappelli, "Effective yet reliable computation of hyperfine coupling constants in solution by a QM/MM approach: Interplay between electrostatics and non-electrostatic effects," J. Chem. Phys. **150**, 124102 (2019).

[85] A. K. Rappe, C. J. Casewit, K. S. Colwell, W. A. Goddard, and W. M. Skiff, "UFF: A full periodic table force field for molecular mechanics and molecular dynamics simulations," J. Am. Chem. Soc. **114**, 10024–10035 (1992).

[86] A. Bondi, "van der Waals volumes and radii," J. Phys. Chem. **68**, 441–451 (1964).

[87] S. A. Kovalenko, R. Schanz, V. M. Farztdinov, H. Hennig, and N. P. Ernsting, "Femtosecond relaxation of photoexcited para-nitroaniline: Solvation, charge transfer, internal conversion and cooling," Chem. Phys. Lett. **323**, 312–322 (2000).

[88] T. Giovannini, R. R. Riso, M. Ambrosetti, A. Puglisi, and C. Cappelli, "Electronic transitions for a fully QM/MM approach based on fluctuating charges and fluctuating dipoles: Linear and corrected linear response regimes," J. Chem. Phys. **151**, 174104 (2019).

[89] T. Giovannini, A. Puglisi, M. Ambrosetti, and C. Cappelli, "Polarizable QM/MM approach with fluctuating charges and fluctuating dipoles: The QM/FQFμ model," J. Chem. Theory Comput. **15**, 2233–2245 (2019).

[90] J. Olsen, P. Jørgensen, H. Koch, A. Balkova, and R. J. Bartlett, "Full configuration–interaction and state of the art correlation calculations on water in a valence double-zeta basis with polarization functions," J. Chem. Phys. **104**, 8007–8015 (1996).

[91] S. D. Folkestad, E. F. Kjønstad, R. H. Myhre, J. H. Andersen, A. Balbi, S. Coriani, T. Giovannini, L. Goletto, T. S. Haugland, A. Hutcheson, I.-M. Høyvik, T. Moitra, A. C. Paul, M. Scavino, A. S. Skeidsvoll, Å. H. Tveten, and H. Koch, Gitlab, 2020, https://gitlab.com/eT-program/eT.

Paper II

Time-dependent coupled-cluster theory for ultrafast transient-absorption spectroscopy

Andreas S. Skeidsvoll ⬤,[1,*] Alice Balbi ⬤,[2,*] and Henrik Koch ⬤[1,2,†]

[1]*Department of Chemistry, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway*
[2]*Scuola Normale Superiore, Piazza dei Cavalieri, 7, IT-56126 Pisa, Italy*

We present a spin-adapted time-dependent coupled-cluster singles and doubles model for the molecular response to a sequence of ultrashort laser pulses. The implementation is used to calculate the electronic response to a valence-exciting pump pulse, and a subsequent core-exciting probe pulse. We assess the accuracy of the integration procedures used in solving the dynamic coupled-cluster equations, in order to find a compromise between computational cost and accuracy. The transient absorption spectrum of lithium fluoride is calculated for various delays of the probe pulse with respect to the pump pulse. We observe that the transient probe absorption oscillates with the pump-probe delay, an effect that is attributed to the interference of states in the pump-induced superposition.

I. INTRODUCTION

Recent advances in the field of ultrafast pulse shaping have enabled the generation of broadband few- to subfemtosecond laser pulses in the near-infrared to vacuum ultraviolet spectral ranges [1–3]. These ultrashort pulses open the possibility to study valence electron dynamics of molecules, on time scales shorter than times characteristic for nuclear dynamics. Also, the generation of intense isolated soft-x-ray free-electron laser pulses with subfemtosecond temporal widths has recently been achieved [4]. This paves the way for attosecond-resolved core-level spectroscopy at high intensities and repetition rates.

Core excitations are typically local to specific atoms, and are sensitive to their electronic environment [5]. The associated attosecond-resolved transient absorption can thus be used to observe superpositions of valence-excited states from the point of view of a specific atomic site, provided that the superposition is of a certain degree of coherence [6]. In the short-pulse limit, the energy-integrated absorption of a core-exciting pulse is indicative of the electronic hole density in the valence region around the nucleus of the specific atom [7,8]. For subfemtosecond pulses outside this limit, the relationship between the pump-induced charge migration and the resultant transient absorption of the probe pulse is more complex. Thus more complete theoretical models are necessary for guiding the pump-probe experiments and for interpreting ensuing results.

Provided that the transient absorption of a probe pulse can be modeled and understood, the valence-level pump and subsequent core-level probe by ultrafast pulses can then be used to investigate the valence electron response of molecules [7,8]. A refined conceptual understanding of this response will shed light on processes occurring in nature, such as photosynthesis and eyesight, and be used for the advance-ment of technological applications, such as photovoltaics and photocatalysis.

Nonperturbative modeling of electron dynamics for ultrafast laser-matter interactions offers certain advantages: the models are applicable for a large range of field intensities [9], and the interaction between a molecule and ultrashort pulses resembles experimental setups in a more natural way.

Electron correlation is often important for a qualitative and quantitative description of many-electron systems. The full configuration interaction (FCI) model is computationally impracticable in most situations [10], and thus we advocate the use of coupled-cluster theory in this paper. Other methods have been used to describe electron dynamics, such as real-time density-functional theory (DFT) [11,12]. However, DFT methods are limited by the accuracy of the exchange-correlation functionals, and thus could lead to misinterpretations [13]. Several implementations of real-time coupled-cluster models have been developed in the past, including approaches based on the time-dependent coupled-cluster (TDCC) equations derived by Koch and Jørgensen [14–17], and approaches based on equation of motion (EOM) theory [9,18–24]. These models offer an accurate description of dynamic correlation, and static correlation in excited states. Needless to say, the coupled-cluster models are also inherently size extensive and intensive [25]. This while keeping the polynomial scaling of the computational costs with respect to system size.

A spin-unrestricted time-dependent coupled-cluster singles and doubles (TDCCSD) model was recently implemented by Pedersen and Kvaal, and used to calculate the absorption spectra of helium and beryllium irradiated by ultrashort pulses at various intensities [26]. Even above the perturbative limit, the TDCCSD spectra show promising correspondence with spectra calculated with time-dependent FCI. The authors also noted that the Lagrangian time-dependent equations have a Hamiltonian structure, well suited for the use of symplectic integrators.

In this work, we will continue the discussion of TDCC models, by presenting a spin-adapted TDCC model of

---

ultrafast transient absorption spectroscopy. Applied to closed-shell molecules interacting with laser pulses within the dipole approximation, this model offers equivalent results as its spin-unrestricted counterparts, with lower computational costs. The reduced cost implies that larger molecules can be studied within this model, making progress towards the accurate modeling of correlated dynamics in interesting photoactive molecules.

This paper is organized as follows. In Sec. II we present the theory underlying the TDCC model and discuss a generalization of the Ehrenfest theorem in this framework. We also describe how absorption spectra are calculated. In Sec. III, we optimize the different parameters used in TDCC calculations, and illustrate this for the LiH molecule. The model is applied to transient absorption of the LiF molecule. Final remarks are given in Sec. IV.

II. THEORY

A. Spin-adapted coupled-cluster method

An accurate account of the electron correlation in molecules is offered by coupled-cluster models, in which the time-independent wave function can be written as

$$|CC\rangle = e^T |HF\rangle, \tag{1}$$

where $|HF\rangle$ is the closed-shell Hartree-Fock reference determinant and T is the spin-adapted cluster operator. The cluster operator is defined as a linear combination of singlet excitation operators τ_μ,

$$T = \sum_{\mu>0} t_\mu \tau_\mu. \tag{2}$$

The expansion coefficients t_μ are referred to as the amplitudes. The operator T is usually truncated at a given level of excitation, for instance, after single excitations gives the coupled-cluster singles (CCS) model, after double excitations gives the coupled-cluster singles and doubles model (CCSD), and so on.

In the Lagrangian formulation of coupled-cluster theory, which satisfies the Hellman-Feynman theorem, the dual state corresponding to the $|CC\rangle$ state is [27]

$$\langle\Lambda| = \left(\langle HF| + \sum_{\nu>0} \bar{t}_\nu \langle\nu|\right) e^{-T}, \tag{3}$$

where the linear expansion coefficients $\bar{t}_\nu$ will be referred to as the (Lagrange) multipliers. The level of excitations is truncated at the same level as the excitations in the cluster operator. We note that the $|CC\rangle$ state and its dual state $\langle\Lambda|$ are biorthonormal, $\langle\Lambda|CC\rangle = 1$.

In this formulation, the expectation values of operators are given as

$$\langle A\rangle = \langle\Lambda|A|CC\rangle$$
$$= \left(\langle HF| + \sum_{\nu>0} \bar{t}_\nu \langle\nu|\right) \bar{A}|HF\rangle, \tag{4}$$

where the similarity transformed operator is defined as

$$\bar{A} = e^{-T} A \, e^T. \tag{5}$$

The amplitudes and multipliers that parametrize the ground state are determined from [28]

$$\langle\mu|\bar{H}|HF\rangle = 0, \tag{6}$$

$$\langle\Lambda|[H, \tau_\mu]|CC\rangle = 0, \tag{7}$$

and the corresponding ground-state energy E_{CC} is given by

$$E_{CC} = \langle\Lambda|H|CC\rangle$$
$$= \langle HF|H|CC\rangle, \tag{8}$$

where we have used Eq. (6) to eliminate the multiplier contribution.

B. Time-dependent coupled-cluster methods

In order to allow for time dependence in the description, the coupled-cluster state is parametrized as [14]

$$|CC(t)\rangle = e^{T(t)}|HF\rangle e^{i\epsilon(t)}, \tag{9}$$

and the corresponding dual state as

$$\langle\Lambda(t)| = \left(\langle HF| + \sum_{\nu>0} \bar{t}_\nu(t)\langle\nu|\right) e^{-T(t)} e^{-i\epsilon(t)}. \tag{10}$$

The amplitudes t_μ and multipliers $\bar{t}_\mu$ now explicitly depend on time, while the excitation operators τ_μ are still time independent. An overall time-dependent phase $\epsilon(t)$ has also been introduced.

The equation describing the time evolution of the amplitudes $t_\mu(t)$ is obtained from the time-dependent Schrödinger equation for the $|CC\rangle$ state, by projecting onto the corresponding excited determinant $\langle\mu|$. This gives the differential equation

$$\frac{dt_\mu(t)}{dt} = -i\langle\mu|\bar{H}(t)|HF\rangle. \tag{11}$$

The equation describing the time evolution of the multipliers $\bar{t}_\mu(t)$ is obtained by projecting the time-dependent Schrödinger equation for the dual state $\langle\Lambda(t)|$ onto the excited determinants $|\nu\rangle$, giving the differential equation

$$\frac{d\bar{t}_\nu(t)}{dt} = i\left(\langle HF| + \sum_{\mu>0} \bar{t}_\mu(t)\langle\mu|\right)[\bar{H}(t), \tau_\nu]|HF\rangle. \tag{12}$$

The equation for the phase $\epsilon(t)$ is determined by projection onto the $|HF\rangle$ state

$$\frac{d\epsilon(t)}{dt} = -\langle HF|\bar{H}(t)|HF\rangle. \tag{13}$$

Detailed derivations can be found in Ref. [14]. In this framework, the time-dependent expectation value of a generic operator $A(t)$ is defined as

$$\langle A(t)\rangle = \langle\Lambda(t)|A(t)|CC(t)\rangle, \tag{14}$$

where $\langle\Lambda(t)|CC(t)\rangle = 1$.

C. Generalized Ehrenfest theorem and conserved quantities in TDCC

For ease of notation, we suppress the explicit time dependence in this section. Ideally, observables calculated in

truncated TDCC should have the same properties as in the untruncated case, in order to give a faithful representation of the physical system. In this context, we derive a generalized Ehrenfest theorem for truncated TDCC (the detailed derivation is given in Appendix A). We obtain the equation

$$
\frac{d}{dt}\langle\Lambda'|A|\text{CC}\rangle = i\langle\Lambda'|H\,e^{T'}P_n e^{-T'}A|\text{CC}\rangle
$$
$$
- i\langle\Lambda'|A\,e^{T}P_n e^{-T}H|\text{CC}\rangle + \langle\Lambda'|\frac{\partial A}{\partial t}|\text{CC}\rangle,
\tag{15}
$$

where the left $\langle\Lambda'|$ state and the right $|\text{CC}\rangle$ state are independent solutions to the projected time-dependent Schrödinger equation. The projection operator P_n of maximum excitation level n is defined as

$$
P_n = |\text{HF}\rangle\langle\text{HF}| + \sum_{\mu>0}^{n}|\mu\rangle\langle\mu|,
\tag{16}
$$

and in untruncated TDCC, $P_n = \mathbb{1}$. From Eq. (15) we can see that, in untruncated TDCC,

$$
\frac{d}{dt}\langle\Lambda'|A|\text{CC}\rangle = i\langle\Lambda'|[H,A]|\text{CC}\rangle + \langle\Lambda'|\frac{\partial A}{\partial t}|\text{CC}\rangle,
\tag{17}
$$

regardless of the initial values of the amplitudes, multipliers, and phases.

In truncated TDCC, the projection operator cannot in general be replaced by the identity operator, and hence Eq. (15) cannot be simplified further. Still, some conservation laws from untruncated TDCC apply under certain constraints: we see from (15) that the Hamiltonian matrix element $\langle\Lambda'|H|\text{CC}\rangle$ is conserved for a time-independent Hamiltonian operator as long as $T' = T$, regardless of the initial values of the multipliers and phases. The overlap matrix element $\langle\Lambda'|\mathbb{1}|\text{CC}\rangle$ is also conserved for $T' = T$, since $\exp(T)P_n\exp(-T)\mathbb{1}|\text{CC}\rangle = |\text{CC}\rangle$ and $\langle\Lambda'|_{T'=T}\mathbb{1}\exp(T)P_n\exp(-T) = \langle\Lambda'|_{T'=T}$. In conclusion, we note the energy and overlap conservation for a time-independent Hamiltonian in untruncated TDCC, and in truncated TDCC for $T' = T$.

D. Interaction with an external electromagnetic field

In the semiclassical approximation, the electronic Hamiltonian for a molecule interacting with an external electromagnetic field can be written as

$$
H(t) = H_0 + V(t),
\tag{18}
$$

where H_0 is the time-independent electronic Hamiltonian and $V(t)$ is the operator describing the interaction with the external field. We choose to express the interaction in the length gauge and dipole approximation, meaning that the electromagnetic field is represented by an electric field,

$$
V(t) = -\mathbf{d}\cdot\boldsymbol{\mathcal{E}}(t),
\tag{19}
$$

where $\mathbf{d}$ is the electric dipole moment operator. Since this operator is a one-electron operator, it can also be expressed in terms of the molecular-orbital (MO) dipole moment integrals d_{pq} and one-electron singlet excitation operators E_{pq},

$$
\mathbf{d} = \sum_{pq}\mathbf{d}_{pq}E_{pq}.
\tag{20}
$$

Since electric fields are additive, the external electric field $\boldsymbol{\mathcal{E}}(t)$ can be written as a linear combination of individual laser pulses,

$$
\boldsymbol{\mathcal{E}}(t) = \sum_n \boldsymbol{\mathcal{E}}_{0n}\cos[\omega_{0n}(t - t_{0n})]f_n(t),
\tag{21}
$$

where $\boldsymbol{\mathcal{E}}_{0n}$ is the peak electric field of pulse n in its polarization direction, ω_{0n} the carrier (angular) frequency and t_{0n} the central time of the pulse, and $f_n(t)$ an envelope function that determines its shape. A commonly used family of envelopes $f_n(t)$, that resemble physical laser intensity profiles, are the Gaussian functions. Since Gaussian functions have infinite support, we choose to set them to zero at a finite number N of root-mean-square (rms) widths σ_n outside the central time, i.e.,

$$
f_n(t) = \begin{cases} e^{-(t-t_{0n})^2/(2\sigma_n^2)}, & a_n \leqslant t \leqslant b_n, \\ 0, & \text{otherwise}, \end{cases}
\tag{22}
$$

where $a_n = t_{0n} - N\sigma_n$ and $b_n = t_{0n} + N\sigma_n$. In addition to resembling physical intensity profiles, a useful feature of Gaussian envelopes is that they give pulses with Gaussian frequency distributions. Hence these pulses can offer a good compromise between temporal precision and spectral narrowness. This is useful for producing temporally precise electronic transitions within the molecule, while keeping the probability of ionization low.

E. Frequency-resolved transient absorption

Following the procedure of [29], the energy absorbed during the interaction with the external electromagnetic field can be given by

$$
\Delta E = \int_{-\infty}^{\infty}\frac{dE(t)}{dt}dt.
\tag{23}
$$

The time derivative of the expectation value of the Hamiltonian in Eq. (18) can be found through Eq. (15):

$$
\frac{dE(t)}{dt} = \frac{d}{dt}\langle\Lambda(t)|H(t)|\text{CC}(t)\rangle
$$
$$
= \langle\Lambda(t)|\frac{\partial H(t)}{\partial t}|\text{CC}(t)\rangle = -\mathbf{d}(t)\cdot\frac{\partial\boldsymbol{\mathcal{E}}(t)}{\partial t},
\tag{24}
$$

where the TDCC dipole moment expectation value is given by

$$
\mathbf{d}(t) = \langle\Lambda(t)|\mathbf{d}|\text{CC}(t)\rangle.
\tag{25}
$$

The energy exchanged between the electromagnetic field and the molecule is thus given by

$$
\Delta E = -\int_{-\infty}^{\infty}\mathbf{d}(t)\cdot\frac{\partial\boldsymbol{\mathcal{E}}(t)}{\partial t}dt.
\tag{26}
$$

Equation (26) can be frequency resolved by inserting the relations between the components $d_i(t)$ and $\mathcal{E}_i(t)$ and their Fourier transforms, $\widetilde{d}_i(\omega)$ and $\widetilde{\mathcal{E}}_i(\omega)$. We use the following convention:

$$
f(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\widetilde{f}(\omega)e^{i\omega t}d\omega,
\tag{27}
$$

$$
\widetilde{f}(\omega) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}f(t)e^{-i\omega t}dt.
\tag{28}
$$

After inserting the relations, the expression

$$\Delta E = \int_0^\infty \omega S(\omega) d\omega \qquad (29)$$

is obtained, where

$$S(\omega) = -2 \operatorname{Im}[\widetilde{\mathbf{d}}(\omega) \cdot \widetilde{\boldsymbol{\mathcal{E}}}^*(\omega)], \quad \omega > 0. \qquad (30)$$

The response function $S(\omega)$ has the opposite sign as in [29], due to different Fourier transform conventions. It represents the absorption per unit frequency at a given frequency, so that positive (negative) $\omega S(\omega)$ equals the amount of energy gained (lost) by the molecule per unit frequency at ω [29].

The TDCC dipole moment $\mathbf{d}(t)$ can be found from Eq. (14)

$$\begin{aligned}
\mathbf{d}(t) &= \sum_{pq} \langle \Lambda(t) | E_{pq} | \mathrm{CC}(t) \rangle \mathbf{d}_{pq} \\
&= \left(\langle \mathrm{HF} | + \sum_{\mu > 0} \bar{t}_\mu(t) \langle \mu | \right) \bar{E}_{pq}(t) | \mathrm{HF} \rangle \mathrm{d}_{pq} \qquad (31) \\
&= \sum_{pq} D_{pq}(t) \mathrm{d}_{pq},
\end{aligned}$$

where $D_{pq}(t)$ is an element of the standard coupled-cluster one-electron density matrix, which can be calculated given the time-dependent amplitudes and multipliers.

F. Initial value problem

In order to calculate the time-dependent amplitudes and multipliers for the system represented by the Hamiltonian in Eq. (18), the system is prepared in the ground state at $t = -T$ (before the interaction). The time-dependent amplitudes and multipliers are then propagated by integration of Eqs. (11) and (12), until $t = T$ (after the interaction). This is done using Runge-Kutta methods (a general introduction to these methods is given in Appendix B). Once the time-dependent amplitudes and multipliers are calculated, they can be used to calculate evenly sampled values of the TDCC dipole moment with Eq. (31).

The main Runge-Kutta method used for integration is the explicit Runge-Kutta (ERK) method known as RK4, and referred to as "the best-known fourth-order four-stage ERK method" in [30]. In many cases, this method gives a good compromise between accuracy and the number of evaluations for each time step.

The performance of two methods in the family of ν-stage 2νth-order implicit Runge-Kutta (IRK) methods, known as Gauss-Legendre methods, is also assessed. An interesting property of these methods is that they are symplectic, meaning that they often perform well with regards to preserving the energy expectation value of noninteracting Hamiltonian systems. The application of these methods to TDCC methods was discussed in greater detail in the work by Pedersen and Kvaal [26]. The Gauss-Legendre methods that will be considered here are the two-stage fourth-order Gauss-Legendre method (GL4) and the three-stage sixth-order Gauss-Legendre method (GL6).

G. Discrete Fourier transformation of TDCC dipole moment and electric field

After the dipole moment and electric field have been calculated in $[-T, T]$, a discrete approximation of $\widetilde{d}_i(\omega)$ and $\widetilde{\mathcal{E}}_i(\omega)$ can be found from doing the discrete Fourier transform of the time series.

Assuming that the finite and discrete time series are sampled from infinitely extending analytic dipole moment and electric-field functions, the time series can equally be represented as the analytic functions modulated by the rectangular window function,

$$f_{w_R}(t) = f(t) w_R(t), \qquad (32)$$

sampled in $[-T, T]$, where the rectangular window function

$$w_R(t) = \begin{cases} 1, & |t| \leqslant T, \\ 0, & \text{otherwise.} \end{cases} \qquad (33)$$

Since the Fourier transform of a windowed function is equal to the convolution of the Fourier transform of the function with the Fourier transform of the window function [31],

$$\widetilde{f_w}(\omega) = \widetilde{f}(\omega) * \widetilde{w}(\omega), \qquad (34)$$

the spectral leakage of the peaks in the finite Fourier spectrum will be related to the Fourier transform of the rectangular window function. In order to reduce the intensity of sidelobes of peaks in the Fourier spectrum [31], the rectangular window can be replaced with a Hann window, by multiplying the sampled values with the Hann function,

$$w_H(t) = \cos^2\left(\frac{\pi t}{T}\right), \qquad (35)$$

before doing the discrete Fourier transform.

III. RESULTS AND DISCUSSION

A. Convergence of LiH pump-probe absorption spectra

In the following, we investigate the convergence properties of the spin-adapted TDCC model of molecular ultrafast pump-probe absorption. The convergence will be assessed with respect to the individual variation of several parameters: the basis set, the size of the time steps, and the integration method. The TDCC method was implemented in the recently released e^T program [32]. This program was used for all reported computations. Unit conversions are done from Hartree atomic units using the 2018 CODATA recommended values [33]. All reported calculations are run on a two-socked node equipped with Intel Xeon-Gold 615 22.1 GHz processors and 1.5 TB of memory.

The higher level coupled-cluster methods scale rapidly with the size of the system, and quickly reach the limits of practicability. Therefore, we have chosen lithium hydride (LiH) for the convergence studies. This serves as an elementary example of a closed-shell molecule with atoms of different core excitation energies. The electronic charge can migrate between the two atoms, making it an interesting case for examination by pump-probe spectroscopy.

The lithium atom is placed at the origin, and the hydrogen atom at $-1.594\,913\,18$ Å along the z axis, corresponding to the experimentally measured equilibrium bond length of LiH

TABLE I. LiH and LiF pump and probe pulse parameters. Gaussian rms width σ, carrier frequency ω_0, peak electric-field strength $|\mathcal{E}_0|$, and the nonperturbative intensity parameter $z_0 = |\mathcal{E}_0|^2/(4\omega_0^3)$ [34] for the carrier waves. A Gaussian rms width of 20 a.u. corresponds to a field strength (intensity) full width at half maximum of 1.139 fs (805.5 as) and 10 a.u to a full width at half maximum of 569.6 as (402.8 as). From the relation $S_0 = |\mathcal{E}_0|^2/Z_0$ (SI units), where Z_0 is the impedance of free space, a peak electric-field strength of 0.01 a.u. corresponds to a peak intensity S_0 of 7.019×10^{12} W/cm^2, and 0.1 a.u. to a peak intensity of 7.019×10^{14} W/cm^2. The perturbation limit can be taken to be the intensity where $z_0 = 1$ for a given carrier frequency [34]. Note that $z_0 \ll 1$ for all pulses, indicating that the interactions also could be described with perturbative approaches.

	σ (a.u.)	ω_0 (eV)	$|\mathcal{E}_0|$ (a.u.)	z_0
LiH pump	20	3.552 47	0.01	1.12×10^{-2}
LiH probe	10	57.6527	0.1	2.63×10^{-4}
LiF pump	20	6.448 01	0.01	1.88×10^{-3}
LiF probe	10	688.018	0.1	1.55×10^{-7}

[35]. Gaussian envelopes are used for the pump and probe pulses, which are polarized in the z direction. The electric fields of each pulse are temporally truncated at eight rms widths σ from the central time, and thus nonzero only inside this interval [see Eq. (22)]. The carrier frequency of the pump pulse corresponds to the first LiH valence excitation energy and the carrier frequency of the probe pulse to the first LiH K-edge excitation energy. These excitation energies are calculated using EOM-CCSD. The core excitations are obtained within the core-valence separation (CVS) approximation [36]. The parameters of the pulses are shown in Table I. As the Gaussian envelopes give the pulses a frequency content distributed around the central frequencies, the pump and probe pulses will induce excitations in the valence and core regions, respectively.

The pump pulse is given a central time of $t = -40$ a.u. and the probe a central time of $t = 0$ a.u. The time-dependent dipole moment and electric field are calculated every 0.1 a.u. in the [-5000 a.u., 5000 a.u.] interval. Since the system remains in the ground state until the onset of the truncated pump pulse—with the ground-state dipole moment—the interaction with the pulses only needs to be calculated in [-200 a.u., 5000 a.u.]. Subsequently, the Hann windowed components of the dipole moment and electric field are discrete Fourier transformed, and the transient absorption is calculated using Eq. (30).

We use the correlation-consistent basis sets of Dunning *et al.* (cc-pVXZ, X = D, T) [37] that are suitable for describing valence correlation effects in molecules. In some of the calculations, the basis sets are augmented by diffuse functions (denoted by aug-) and/or functions describing core correlation (denoted by C) [38]. From now on, we will use a C in round brackets to indicate that core correlation functions are added to the basis set of the heaviest atom in the molecule.

The individual variation of the calculation parameters is done with respect to a common reference: TDCCSD/aug-cc-p(C)VDZ, and integrated with RK4 with 0.005 a.u. time steps. This basis set gives two occupied and 34 virtual
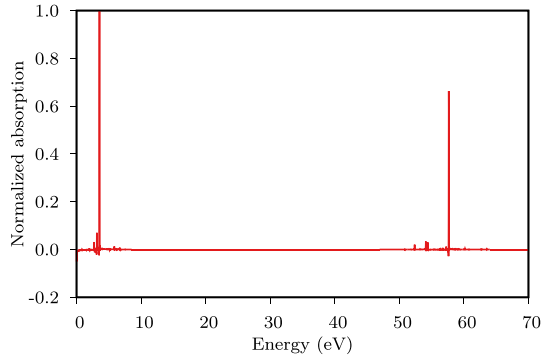

FIG. 1. Normalized reference LiH pump-probe absorption, $S'(\omega)$, as a function of energy. The time-dependent dipole moment is calculated using TDCCSD/aug-cc-p(C)VDZ, and integrated with RK4 with 0.005 a.u. time steps.

MOs, and hence 4828 time-dependent parameters. The reported calculation uses around 1 GB of memory and 0.12 s wall time per time step on eight cores, and 0.15 s per time step on four cores. The unnormalized reference absorption $S_{\mathrm{ref}}(\omega)$ is used to calculate the normalization factor

$$\mathcal{N}_{\mathrm{ref}} = \frac{1}{\max_\omega |S_{\mathrm{ref}}(\omega)|}. \tag{36}$$

This factor is used to normalize all the absorption spectra of the following LiH calculations, by means of

$$S'(\omega) = \mathcal{N}_{\mathrm{ref}} S(\omega), \tag{37}$$

where $S(\omega)$ is calculated with the parameters in question. The normalized deviation of $S'(\omega)$ from a more accurate result $S'_{\mathrm{acc}}(\omega)$ is calculated as

$$D'(\omega) = |S'(\omega) - S'_{\mathrm{acc}}(\omega)|. \tag{38}$$

The reference absorption spectrum, normalized according to Eq. (37), is shown in Fig. 1. We observe absorption in two energy regions: one corresponding to the valence-exciting pump pulse and the other to the core-exciting probe pulse.

1. TDCCS and TDCCSD

In Fig. 2, the normalized reference TDCCSD spectrum is shown together with the normalized time-dependent CCS (TDCCS) spectrum. The two spectra display substantial differences in intensities and positions of the peaks in both the pump and the probe absorption regions. Since TDCCSD includes double excitations, while TDCCS does not, this demonstrates that higher-order excitations are needed to obtain qualitatively correct results for the LiH model system.

2. Basis set

In Fig. 3, the normalized reference spectrum is shown together with normalized spectra calculated using cc-pVDZ, cc-p(C)VDZ, and aug-cc-pVDZ. The inclusion of diffuse functions in the basis sets seems important for representing the dynamics properly. Increasing the basis set from cc-pVDZ to aug-cc-pVDZ shifts the peaks in both the pump and the probe
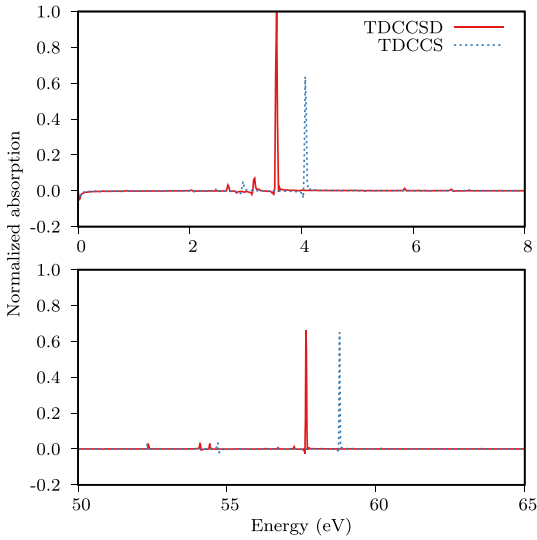
FIG. 2. Normalized TDCCSD and TDCCS LiH pump and probe absorption, $S'(\omega)$, as a function of energy. Time-dependent dipole moments are calculated using aug-cc-p(C)VDZ, and integrated with RK4 with 0.005 a.u. time steps.
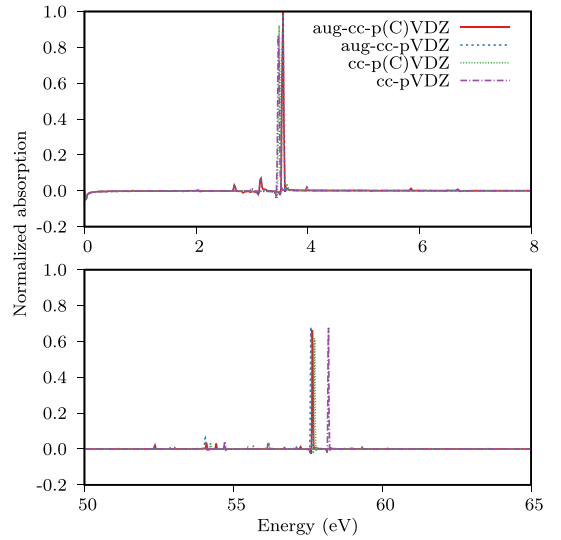

FIG. 3. Normalized aug-cc-p(C)VDZ, aug-cc-pVDZ, cc-p(C)VDZ, and cc-pVDZ LiH pump and probe absorption, $S'(\omega)$, as a function of energy. Time-dependent dipole moments are calculated using TDCCSD, and integrated with RK4 with 0.005 a.u. time steps.

absorption regions. This is consistent with the concept of the pump pulse forcing electrons to the outer valence regions of the molecule, which is better represented with diffuse functions.

Furthermore, comparing cc-p(C)VDZ and cc-pVDZ spectra in Fig. 3, we see the importance of the added core correlation functions. As expected, they cause a substantial shift in the probe absorption peaks, while they are not important for the pump absorption.

We also performed calculations with cc-pVTZ, cc-p(C)VTZ, aug-cc-pVTZ, and aug-cc-p(C)VTZ basis sets. Note that, for the aug-cc-pVTZ and aug-cc-p(C)VTZ spectra, the time-dependent dipole moments are only calculated in the [−2500 a.u., 2500 a.u.] interval, in order to reduce computational time. Thus these spectra have a lower resolution than the others. The normalized spectra are shown together with the normalized reference spectrum in Fig. 4. Here we observe that triple zeta functions change the position of the peaks in the probe absorption region. This indicates that basis sets larger than aug-cc-p(C)VDZ should be used if precise peak positions are required, bringing about a substantial increase in the computational costs. The aug-cc-p(C)VDZ basis set is used as the reference for the other LiH calculations, as the larger basis sets are too computationally expensive for practical purposes.

Note that the pulses are not strong enough to induce considerable multiphoton absorption (see Table I). The electrons should thus primarily be confined to low angular momentum bound states, which are fairly well described with the aforementioned basis sets. At higher intensities, the results obtained with these basis sets should deviate further from the complete basis set limit, as the representation of Rydberg

states and the continuum will be more important [39]. The results can then be improved by adding suitable functions to the basis set, for instance, Gaussians optimized for the representation of the continuum [39,40].

3. Integration

We calculated normalized spectra for 0.125 a.u., 0.025 a.u., and 0.001 a.u. time steps. The deviations from the 0.001 a.u. time step are calculated according to Eq. (38). The results are shown in Fig. 5. The deviations decrease with the time step size, indicating that the spectra approach a time step limit.

We further calculated normalized spectra with GL4 and GL6. The deviations of the RK4 (reference) and GL4 spectra from the GL6 spectrum are shown in Fig. 6. Although the TDCC equations have a Hamiltonian structure, the use of symplectic integrators does not seem to be necessary to calculate accurate spectra for this system, with the applied field strength. As the three integration methods give comparable results, we will use RK4 for the other calculations, as this generally requires fewer evaluations of the TDCC equations per time step.

B. LiF transient absorption

In this section, variations in molecular absorption caused by ultrafast charge migration are modeled in the described pump-probe framework. We consider the lithium fluoride (LiF) molecule, where the fluorine atom is placed at the origin and the lithium atom at −1.563 864 13 Å along the z axis. This corresponds to the experimentally measured equilibrium bond length of LiF [35]. In order to classify some of the transitions
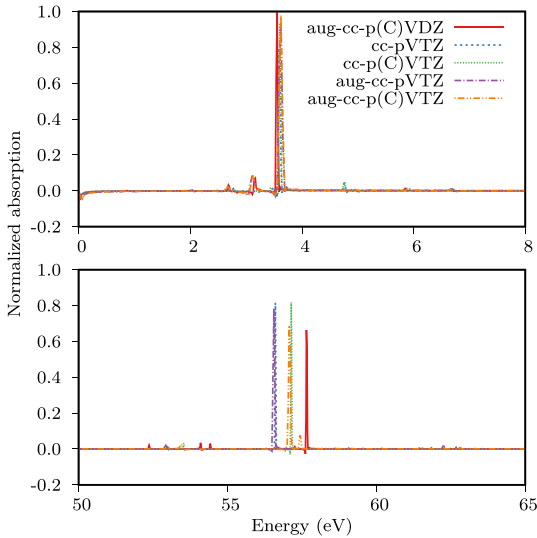
FIG. 4. Normalized aug-cc-p(C)VDZ, cc-pVTZ, cc-p(C)VTZ, aug-cc-pVTZ, and aug-cc-p(C)VTZ LiH pump-probe absorption, $S'(\omega)$, as a function of energy. Time-dependent dipole moments are calculated using TDCCSD, and integrated with RK4 with 0.005 a.u. time steps.

involved in the molecular absorption, the first eight valence-excited and the first eight core-excited states are calculated using EOM-CCSD/aug-cc-p(C)VDZ. The core excitations are obtained within the CVS approximation. The molecular term symbols and excitation energies are given in Table II.

In the TDCC calculations, all probe pulses are z polarized, and have carrier frequencies corresponding to the first LiF core excitation energy (see Table II). Central times are chosen to be 0 a.u., to minimize the effect of the windowing on
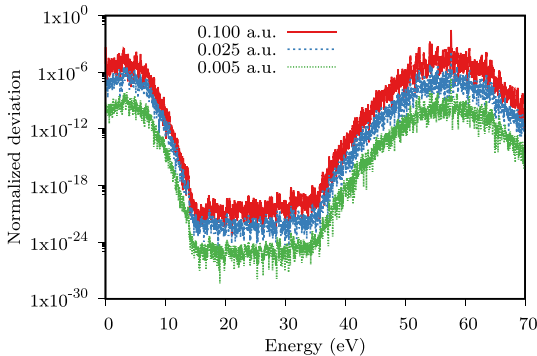
FIG. 5. LiH pump-probe absorption. Normalized deviation of the 0.100 a.u., 0.025 a.u., and 0.005 a.u. time step spectra from the 0.001 a.u. time step spectrum, $D'(\omega)$, as a function of energy. Time-dependent dipole moments are calculated using TDCCSD/aug-cc-p(C)VDZ, and integrated with RK4.
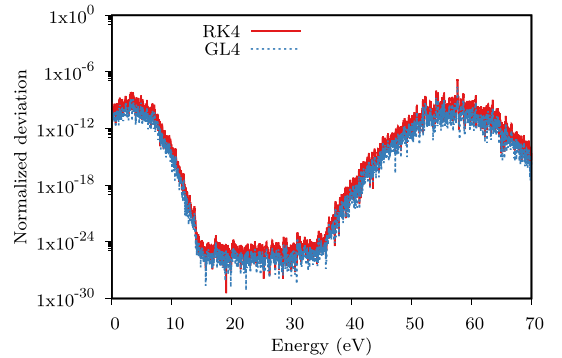
FIG. 6. LiH pump-probe absorption. Normalized deviation of the RK4 and GL4 spectra from the GL6 spectrum, $D'(\omega)$, as a function of energy. Time-dependent dipole moments are calculated using TDCCSD/aug-cc-p(C)VDZ, and integrated with 0.005 a.u. time steps.

the probe absorption. The pump pulses are also z polarized, and have carrier frequencies corresponding to the first LiF valence excitation energy (see Table II). The pump pulses have different central times with respect to the probe pulses, corresponding to probe delays from 0 a.u. to 240 a.u., in increments of 5 a.u. Other parameters of the pump and probe pulses are given in Table I. As for the LiH calculations, the electric fields of each pulse are temporally truncated at eight rms widths σ from the central time, and thus nonzero only inside this interval [see Eq. (22)].

The parameters used for the LiH reference calculation offered a compromise between computational cost and accuracy. For pragmatic reasons, we also use the parameters for all LiF calculations. The calculations in this section are thus done using TDCCSD/aug-cc-p(C)VDZ, and integrated with RK4 with 0.005 a.u. time steps. The basis set gives six occupied and 44 virtual MOs, and hence 70488 time-dependent parameters. The time-dependent dipole moments and electric fields are calculated every 0.1 a.u. in the [−5000 a.u., 5000 a.u.] interval, where the external field interactions are only calculated after the onset of the temporally truncated pump pulses. The

TABLE II. Molecular term symbols and ground-state excitation energies ΔE of some excited states of LiF, calculated with the EOM-CCSD method. Valence-excited states are denoted by a subscript v. Core-excited states, calculated within the CVS approximation, are denoted by a subscript c.

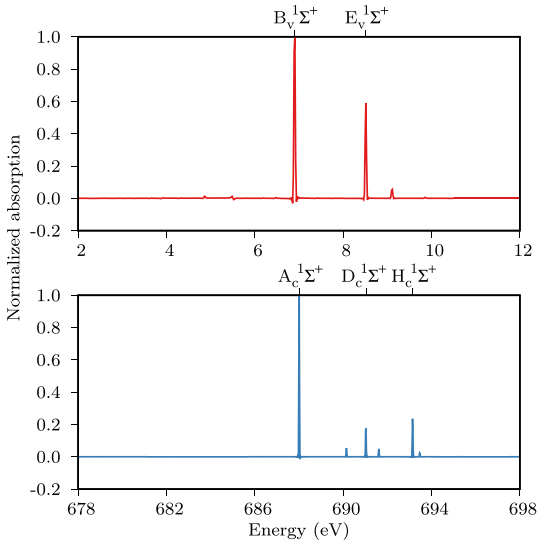| State	ΔE (eV)	State	ΔE (eV)
$A_v\,^1\Pi$	6.448 01	$A_c\,^1\Sigma^+$	688.018
$B_v\,^1\Sigma^+$	6.899 82	$B_c\,^1\Pi$	689.462
$C_v\,^1\Delta$	8.104 63	$C_c\,^1\Sigma^+$	690.159
$D_v\,^1\Sigma^-$	8.140 74	$D_c\,^1\Sigma^+$	691.039
$E_v\,^1\Sigma^+$	8.511 16	$E_c\,^1\Pi$	691.435
$F_v\,^1\Pi$	8.589 43	$F_c\,^1\Sigma^+$	691.625
$G_v\,^1\Pi$	8.625 89	$G_c\,^1\Pi$	692.917
$H_v\,^1\Sigma^+$	9.106 55	$H_c\,^1\Sigma^+$	693.154

FIG. 7. Normalized LiF pump and probe absorption, $S'_{\text{pump}}(\omega)$ (top) and $S'_{\text{probe}}(\omega)$ (bottom), as a function of energy. The most dominant peaks are identified with ground-state transitions to EOM-CCSD valence- and core-excited states. Time-dependent dipole moments are calculated with TDCCSD/aug-cc-p(C)VDZ, integrated with RK4 with 0.005 a.u. time steps.

reported calculations use around 2 GB of memory and 0.41 s wall time per time step on 16 cores, and 0.50 s per time step on eight cores.

In order to assess the relative occupation of the states in the pump-induced superposition [see Eq. (29)], the normalized absorption of the pump pulse, centered at 0 a.u., is calculated using

$$S'_{\text{pump}}(\omega) = \mathcal{N}_{\text{pump}} S_{\text{pump}}(\omega), \tag{39}$$

where

$$\mathcal{N}_{\text{pump}} = \frac{1}{\max_{\omega} |S_{\text{pump}}(\omega)|}. \tag{40}$$

An analogous procedure is used to obtain the normalized probe spectrum $S'_{\text{probe}}(\omega)$.

The normalized absorption of the pump pulse and of the probe pulse are plotted in Fig. 7, where the most dominant absorption peaks are identified using the calculated EOM-CCSD states (see Table II). The small pump absorption peaks that lie below the ground-state valence excitation energy gap are presumably caused by two-photon absorption. The positions of the other visible peaks in the two spectra fit well with single-photon EOM-CCSD transitions allowed by symmetry.

The pump-probe absorption $S(\omega, \tau)$ is calculated as a function of the energy, ω, and the delay of the probe pulse with respect to the pump pulse, τ. In order to directly assess the change in absorption caused by the interaction with the pump pulse, the normalized transient absorption

$$\Delta S'(\omega, \tau) = \mathcal{N}_{\text{probe}} \Delta S(\omega, \tau)$$
$$= \mathcal{N}_{\text{probe}}[S(\omega, \tau) - S_{\text{probe}}(\omega)] \tag{41}$$

is calculated for all delays, where $\mathcal{N}_{\text{probe}}$ is the normalization factor for the probe spectrum. The normalized transient absorption in the probe absorption region is shown in Fig. 8. The spectrum features several constant energy peaks that oscillate with the pump-probe delay. The five peaks that oscillate the
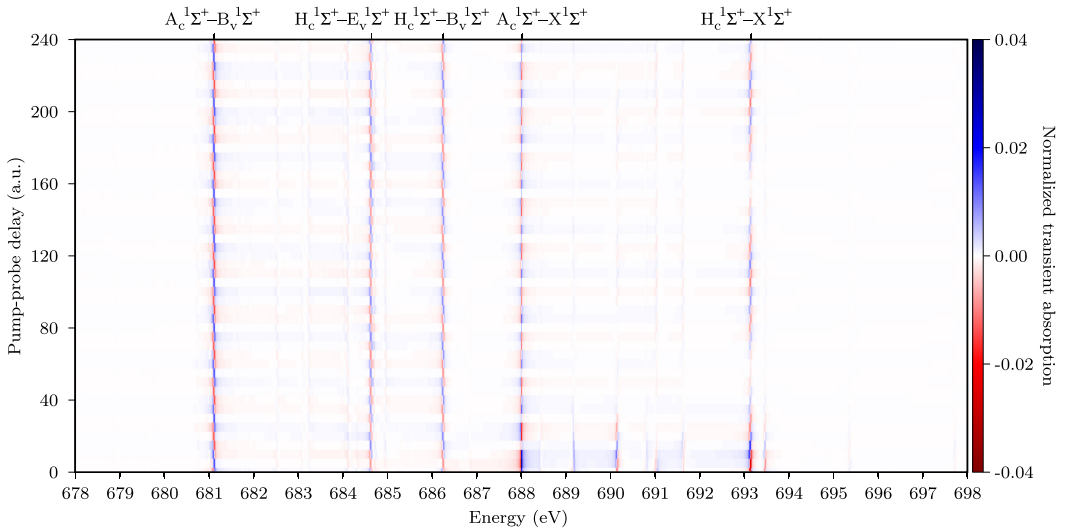
FIG. 8. Normalized LiF transient absorption $\Delta S'(\omega, \tau)$, as a function of energy and pump-probe delay. The five peaks oscillating with the largest amplitude are identified with EOM-CCSD transitions. Time-dependent dipole moments are calculated using TDCCSD/aug-cc-p(C)VDZ basis set, integrated with RK4 with 0.005 a.u. time steps.
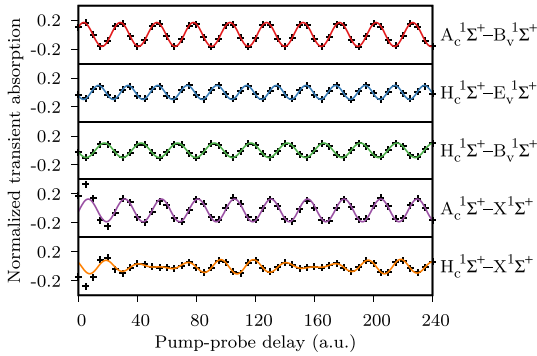
FIG. 9. Normalized LiF transient absorption $\Delta S'(\omega, \tau)$ (black crosses) as a function of pump-probe delay, given at the discrete Fourier transform energies closest to the energies of the transitions shown to the right. The colored functions in the four topmost panels are found from least-squares fitting $A \sin(\omega_A t + \phi_A) + C$, with fixed values of ω_A, to the absorption, in the domain [40 a.u., 240 a.u.]. The values of ω_A are 6.899 82 eV (red) 8.511 16 eV (blue), 6.899 82 eV (green), and 6.899 82 eV (purple). The orange function in the bottom panel is found from least-squares fitting $A \sin(\omega_A t + \phi_A) + B \sin(\omega_B t + \phi_B) + C$, with $\omega_A =$ 6.899 82 eV and $\omega_B = 8.511$ 16 eV, to the absorption, in the domain [40 a.u., 240 a.u.].

most with respect to the pump-probe delay are identified using the states in Table II. Note that, for pump-probe delays shorter than about 40 a.u., the oscillations of some of the peaks are rapidly damped as a function of increasing delays. This effect can be attributed to the decreasing overlap between the pump and probe pulses. For pump-probe delays longer than about 40 a.u., where the overlap of the pulses can be neglected, the damping of the oscillations is also negligible.

We note that the excitation by the pump pulse enables new transitions in the probe absorption region. An illustrative example is the oscillating peak at around 681.1 eV in Fig. 8. The energy corresponding to this peak is lower than the lowest ground-state core excitation energy of 688.018 eV. This peak is identified as the $A_c\,^1\Sigma^+ - B_v\,^1\Sigma^+$ transition. Its occurrence indicates that the pump has generated an electronic hole in a previously occupied region of the molecule, allowing a lower-energy core excitation to take place.

In Fig. 9, the normalized transient absorption of the five peaks identified in Fig. 8 are plotted at the nearest discrete Fourier-transform energies. Two of these peaks describe transitions involving the $A_c\,^1\Sigma^+$ state. Beyond the pump-probe overlap region, the oscillations of these peaks correlate with the quantum interference of the two probed states, as expected for the ultrafast high-energy probing of two states in a coherent superposition [6]. This since both oscillations can be fitted with sinusoids with the frequency corresponding to the $B_v\,^1\Sigma^+$ and $X\,^1\Sigma^+$ energy difference.

Three peaks in Fig. 9 correspond to transitions involving the $H_c\,^1\Sigma^+$ state. The oscillation of the $H_c\,^1\Sigma^+ - B_v\,^1\Sigma^+$ peak correlates well with the quantum interference of the $B_v\,^1\Sigma^+$ and $X\,^1\Sigma^+$ states, as the oscillations are well fitted with a sinusoid with the frequency corresponding to the energy

difference of these two states. Similarly, the oscillation of the $H_c\,^1\Sigma^+ - E_v\,^1\Sigma^+$ peak correlates with the quantum interference of the $E_v\,^1\Sigma^+$ and $X\,^1\Sigma^+$ states. Note that the oscillations of the two peaks are slightly phase shifted with respect to each other, an effect that may be caused by the difference in spectral phase of the two corresponding frequencies in the probe pulse.

The linear combination of two sinusoids is needed to give a good fit with the oscillation of the $H_c\,^1\Sigma^+ - X\,^1\Sigma^+$ peak: one corresponding to the $B_v\,^1\Sigma^+$ and $X\,^1\Sigma^+$ energy difference and the other corresponding to the $E_v\,^1\Sigma^+$ and $X\,^1\Sigma^+$ energy difference. Hence the ground state $X\,^1\Sigma^+$ seems to have a similar probability of interfering with the $B_v\,^1\Sigma^+$ and $E_v\,^1\Sigma^+$ states. This is reasonable, considering that most of the population will be left in the ground state after the interaction with the pump pulse.

IV. CONCLUSION

In this work, a time-dependent coupled-cluster model of ultrafast pump-probe absorption spectroscopy has been presented. First, we investigated the convergence of LiH pump-probe absorption spectra with respect to different calculation parameters. The deviations related to the integration parameters (integration method and time step size) were small in comparison to other parameter-dependent deviations. As the computational costs scaled linearly with the time step size, we chose a time step size that gave a small deviation, 0.005 a.u. In future works, calculations on larger systems can be facilitated by the use of larger and adaptive time steps, as the maximum normalized deviation of the absorption calculated with 0.025 a.u. time steps was only on the order of 1×10^{-4}. The use of symplectic integrators did not seem to be necessary; hence RK4 was used. Changes in the basis set had a big impact on the results. As the computational cost scales steeply with respect to the basis set, TDCCSD/aug-cc-p(C)VDZ was chosen as a compromise between accuracy and computational cost.

After using the time-dependent coupled-cluster model to assess the convergence of LiH spectra, we used the model to calculate the ultrafast transient absorption in LiF, using the same parameters. The transient absorption displayed peaks that oscillate with respect to pump-probe delay, and the oscillation frequencies were correlated with the quantum interference of different states in the pump-induced superposition.

Note that nuclear motion, which has been neglected in the model, will cause broadening of the spectral peaks [41]. A natural next step would be to include the nuclear motion to the model, which for instance can be done using the approach in [42].

APPENDIX A: DERIVATION OF GENERALIZED EHRENFEST THEOREM IN TRUNCATED TDCC

For ease of notation, the time dependence is not written explicitly in this section. The derivation of Eq. (15) in truncated TDCC is given here. It makes use of the identity resolution

$$\mathbb{1} = |\text{HF}\rangle\langle\text{HF}| + \sum_{\mu>0} |\mu\rangle\langle\mu|, \tag{A1}$$

where the summation is over all the excited determinants. Sums that are restricted to the excited determinants in the projection space will be denoted by the upper summation limit n.

Consider a generic operator A with no parametric time dependence and two independent solutions to the projected time-dependent Schrödinger equation, $|\text{CC}\rangle$ and $\langle\Lambda'|$. The time derivative of the matrix element $\langle\Lambda'|A|\text{CC}\rangle$ is

$$\frac{d}{dt}\langle\Lambda'|A|\text{CC}\rangle = \left(\frac{d}{dt}\langle\Lambda'|\right)A|\text{CC}\rangle + \langle\Lambda'|\frac{\partial A}{\partial t}|\text{CC}\rangle$$
$$+ \langle\Lambda'|A\left(\frac{d}{dt}|\text{CC}\rangle\right). \tag{A2}$$

Equations (11) and (13) can be used to rewrite the term containing the time derivative of the $|\text{CC}\rangle$ state,

$$\langle\Lambda'|A\left(\frac{d}{dt}|\text{CC}\rangle\right) = \sum_{\mu>0}^{n}\langle\Lambda'|A\tau_\mu|\text{CC}\rangle\frac{dt_\mu}{dt} + i\langle\Lambda'|A|\text{CC}\rangle\frac{d\epsilon}{dt}$$
$$= -i\langle\Lambda'|A\,e^T P_n e^{-T} H|\text{CC}\rangle. \tag{A3}$$

Equations (11), (12), and (13) can be used to rewrite the term containing the time derivative of the $\langle\Lambda'|$ state,

$$\left(\frac{d}{dt}\langle\Lambda'|\right)A|\text{CC}\rangle = \sum_{\mu>0}^{n}\frac{d\bar{t}'_\mu}{dt}\langle\mu|e^{-T'}e^{-i\epsilon'}A|\text{CC}\rangle$$
$$- \sum_{\mu>0}^{n}\langle\Lambda'|\tau_\mu A|\text{CC}\rangle\frac{dt'_\mu}{dt} - i\langle\Lambda'|A|\text{CC}\rangle\frac{d\epsilon'}{dt}$$
$$= \sum_{\mu>0}^{n}i\langle\Lambda'|H\,e^{T'}|\mu\rangle\langle\mu|e^{-T'}A|\text{CC}\rangle$$
$$- \sum_{\mu>0}^{n}i\langle\Lambda'|e^{T'}\tau_\mu\bar{H}'|\text{HF}\rangle\langle\mu|e^{-T'}A|\text{CC}\rangle$$
$$+ \sum_{\mu>0}^{n}i\langle\Lambda'|\tau_\mu A|\text{CC}\rangle\langle\mu|\bar{H}'|\text{HF}\rangle$$
$$+ i\langle\Lambda'|A|\text{CC}\rangle\langle\text{HF}|\bar{H}'|\text{HF}\rangle. \tag{A4}$$

The right-hand side of Eq. (A1) is inserted between τ_μ and $\bar{H}'$ in the second term, giving

$$\left(\frac{d}{dt}\langle\Lambda'|\right)A|\text{CC}\rangle$$
$$= \sum_{\mu>0}^{n}i\langle\Lambda'|H\,e^{T'}|\mu\rangle\langle\mu|e^{-T'}A|\text{CC}\rangle$$
$$- \sum_{\mu>0}^{n}i\langle\Lambda'|e^{T'}|\mu\rangle\langle\text{HF}|\bar{H}'|\text{HF}\rangle\langle\mu|e^{-T'}A|\text{CC}\rangle$$
$$- \sum_{\mu>0}^{n}\sum_{\nu>0}i\langle\Lambda'|e^{T'}\tau_\mu|\nu\rangle\langle\nu|\bar{H}'|\text{HF}\rangle\langle\mu|e^{-T'}A|\text{CC}\rangle$$
$$+ \sum_{\mu>0}^{n}i\langle\Lambda'|\tau_\mu A|\text{CC}\rangle\langle\mu|\bar{H}'|\text{HF}\rangle$$
$$+ i\langle\Lambda'|A|\text{CC}\rangle\langle\text{HF}|\bar{H}'|\text{HF}\rangle$$
$$= \sum_{\mu>0}^{n}i\langle\Lambda'|H\,e^{T'}|\mu\rangle\langle\mu|e^{-T'}A|\text{CC}\rangle$$
$$- \sum_{\mu>0}^{n}i\langle\Lambda'|e^{T'}|\mu\rangle\langle\mu|e^{-T'}A|\text{CC}\rangle\langle\text{HF}|\bar{H}'|\text{HF}\rangle$$
$$- \sum_{\nu>0}^{n}\sum_{\mu>0}i\langle\Lambda'|e^{T'}\tau_\nu|\mu\rangle\langle\mu|e^{-T'}A|\text{CC}\rangle\langle\nu|\bar{H}'|\text{HF}\rangle$$
$$+ \sum_{\mu>0}^{n}i\langle\Lambda'|\tau_\mu A|\text{CC}\rangle\langle\mu|\bar{H}'|\text{HF}\rangle$$
$$+ i\langle\Lambda'|A|\text{CC}\rangle\langle\text{HF}|\bar{H}'|\text{HF}\rangle. \tag{A5}$$

The factors $\sum_{\mu>0}|\mu\rangle\langle\mu|$ in the second and third terms are replaced by using Eq. (A1), with $|\text{HF}\rangle\langle\text{HF}|$ subtracted from both sides of the equation, giving

$$\left(\frac{d}{dt}\langle\Lambda'|\right)A|\text{CC}\rangle$$
$$= \sum_{\mu>0}^{n}i\langle\Lambda'|He^{T'}|\mu\rangle\langle\mu|e^{-T'}A|\text{CC}\rangle$$
$$+ i\langle\Lambda'|e^{T'}|\text{HF}\rangle\langle\text{HF}|e^{-T'}A|\text{CC}\rangle\langle\text{HF}|\bar{H}'|\text{HF}\rangle$$
$$+ \sum_{\nu>0}^{n}i\langle\Lambda'|e^{T'}|\nu\rangle\langle\text{HF}|e^{-T'}A|\text{CC}\rangle\langle\nu|\bar{H}'|\text{HF}\rangle$$
$$= \sum_{\mu>0}^{n}i\langle\Lambda'|H\,e^{T'}|\mu\rangle\langle\mu|e^{-T'}A|\text{CC}\rangle$$
$$+ i\langle\Lambda'|e^{T'}|\text{HF}\rangle\langle\text{HF}|\bar{H}'|\text{HF}\rangle\langle\text{HF}|e^{-T'}A|\text{CC}\rangle$$
$$+ \sum_{\nu>0}^{n}i\langle\Lambda'|e^{T'}|\nu\rangle\langle\nu|\bar{H}'|\text{HF}\rangle\langle\text{HF}|e^{-T'}A|\text{CC}\rangle$$
$$= i\langle\Lambda'|H\,e^{T'}P_n e^{-T'}A|\text{CC}\rangle, \tag{A6}$$

where the definition of P_n is given in Eq. (16). Equations (A3) and (A6) are inserted into Eq. (A2), giving the desired result:

$$\frac{d}{dt}\langle\Lambda'|A|\text{CC}\rangle = i\langle\Lambda'|H\,e^{T'}P_n e^{-T'}A|\text{CC}\rangle$$
$$- i\langle\Lambda'|A\,e^{T}P_n e^{-T}H|\text{CC}\rangle$$
$$+ \langle\Lambda'|\frac{\partial A}{\partial t}|\text{CC}\rangle. \tag{A7}$$

APPENDIX B: RUNGE-KUTTA METHODS

The commonly used one-step integration methods known as Runge-Kutta methods are introduced below in the notation of [30].

Given the following Cauchy problem:

$$\frac{d\mathbf{y}(t)}{dt} = \mathbf{f}(t, \mathbf{y}(t)), \quad t \geqslant t_0, \ \mathbf{y}(t_0) = \mathbf{y}_0, \tag{B1}$$

we can find a numerical approximation of the solution $\mathbf{y}(t)$ by the use of a ν-stage Runge-Kutta method, which can be written in the form

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h\sum_{j=1}^{\nu} b_j \mathbf{f}(t_n + c_j h, \xi_j), \tag{B2}$$

where

$$\xi_j = \mathbf{y}_n + h\sum_{i=1}^{\nu} a_{ji}\mathbf{f}(t_n + c_i h, \xi_i), \quad j = 1, \dots, \nu. \tag{B3}$$

Here, a_{ji}, b_j, and c_j are method specific coefficients, where a_{ji} and c_j need to satisfy the condition

$$\sum_{j=1}^{\nu} a_{ji} = c_j, \quad j = 1, \dots, \nu \tag{B4}$$

to obtain nontrivial orders of integration. In explicit Runge-Kutta (ERK) methods, the matrix $A = (a_{ji})_{j,i=1,\dots,\nu}$ is strictly lower triangular. In these methods, ξ_j are explicitly given as a function of $\xi_{j-1}, \dots, \xi_1$.

In the cases where the matrix A is not strictly lower triangular, ξ_j may also depend on $\xi_j, \dots, \xi_\nu$, which in practice means that a system of equations have to be solved at each time step. These methods are known as implicit Runge-Kutta (IRK) methods, and in many cases offer greater stability than their explicit counterparts. Since IRK methods involve the solution of a set of equations at each time step, it is hard to give an *a priori* estimate of the number of function evaluations needed at each time step. This number is usually higher than for ERK methods, leading in general to higher computational costs.

[1] M. T. Hassan, T. T. Luu, A. Moulet, O. Raskazovskaya, P. Zhokhov, M. Garg, N. Karpowicz, A. M. Zheltikov, V. Pervak, F. Krausz, and E. Goulielmakis, Nature (London) **530**, 66 (2016).

[2] M. Galli, V. Wanie, D. P. Lopes, E. P. Månsson, A. Trabattoni, L. Colaizzi, K. Saraswathula, A. Cartella, F. Frassetto, L. Poletto, F. Légaré, S. Stagira, M. Nisoli, R. M. Vázquez, R. Osellame, and F. Calegari, Opt. Lett. **44**, 1308 (2019).

[3] D. Fabris, T. Witting, W. A. Okell, D. J. Walke, P. Matia-Hernando, J. Henkel, T. R. Barillot, M. Lein, J. P. Marangos, and J. W. G. Tisch, Nat. Photon. **9**, 383 (2015).

[4] J. Duris, S. Li, T. Driver, E. G. Champenois, J. P. MacArthur, A. A. Lutman, Z. Zhang, P. Rosenberger, J. W. Aldrich, R. Coffee, G. Coslovich, F.-J. Decker, J. M. Glownia, G. Hartmann, W. Helml, A. Kamalov, J. Knurr, J. Krzywinski, M.-F. Lin, J. P. Marangos, M. Nantel, A. Natan, J. T. O'Neal, N. Shivaram, P. Walter, A. L. Wang, J. J. Welch, T. J. A. Wolf, J. Z. Xu, M. F. Kling, P. H. Bucksbaum, A. Zholents, Z. Huang, J. P. Cryan, and A. Marinelli, Nat. Photon. **14**, 30 (2020).

[5] T. J. A. Wolf, R. H. Myhre, J. P. Cryan, S. Coriani, R. J. Squibb, A. Battistoni, N. Berrah, C. Bostedt, P. Bucksbaum, G. Coslovich, R. Feifel, K. J. Gaffney, J. Grilj, T. J. Martinez, S. Miyabe, S. P. Moeller, M. Mucke, A. Natan, R. Obaid, T. Osipov, O. Plekan, S. Wang, H. Koch, and M. Gühr, Nat. Commun. **8**, 29 (2017).

[6] E. Goulielmakis, Z.-H. Loh, A. Wirth, R. Santra, N. Rohringer, V. S. Yakovlev, S. Zherebtsov, T. Pfeifer, A. M. Azzeer, M. F. Kling, S. R. Leone, and F. Krausz, Nature (London) **466**, 739 (2010).

[7] A. D. Dutoi, K. Gokhberg, and L. S. Cederbaum, Phys. Rev. A **88**, 013419 (2013).

[8] A. D. Dutoi and L. S. Cederbaum, Phys. Rev. A **90**, 023414 (2014).

[9] D. R. Nascimento and A. E. DePrince, J. Chem. Theory Comput. **12**, 5834 (2016).

[10] J. Olsen, P. Jørgensen, and J. Simons, Chem. Phys. Lett. **169**, 463 (1990).

[11] P. Hessler, N. T. Maitra, and K. Burke, J. Chem. Phys. **117**, 72 (2002).

[12] N. T. Maitra, J. Chem. Phys. **144**, 220901 (2016).

[13] A. Bruner, S. Hernandez, F. Mauger, P. M. Abanador, D. J. LaMaster, M. B. Gaarde, K. J. Schafer, and K. Lopata, J. Phys. Chem. Lett. **8**, 3991 (2017).

[14] H. Koch and P. Jørgensen, J. Chem. Phys. **93**, 3333 (1990).

[15] S. Kvaal, J. Chem. Phys. **136**, 194109 (2012).

[16] D. A. Pigg, G. Hagen, H. Nam, and T. Papenbrock, Phys. Rev. C **86**, 014308 (2012).

[17] H. E. Kristiansen, Ø. S. Schøyen, S. Kvaal, and T. B. Pedersen, J. Chem. Phys. **152**, 71102 (2020).

[18] J. A. Sonk and H. B. Schlegel, J. Phys. Chem. A **115**, 11832 (2011).

[19] J. A. Sonk, M. Caricato, and H. B. Schlegel, J. Phys. Chem. A **115**, 4678 (2011).

[20] E. Luppi and M. Head-Gordon, Mol. Phys. **110**, 909 (2012).

[21] D. R. Nascimento and A. E. DePrince, J. Phys. Chem. Lett. **8**, 2951 (2017).

[22] L. N. Koulias, D. B. Williams-Young, D. R. Nascimento, A. E. DePrince, and X. Li, J. Chem. Theory Comput. **15**, 6617 (2019).

[23] D. R. Nascimento and A. E. DePrince, J. Chem. Phys. **151**, 204107 (2019).

[24] Y. C. Park, A. Perera, and R. J. Bartlett, J. Chem. Phys. **151**, 164117 (2019).

[25] H. Koch, H. J. A. Jensen, P. Jørgensen, and T. Helgaker, J. Chem. Phys. **93**, 3345 (1990).

[26] T. B. Pedersen and S. Kvaal, J. Chem. Phys. **150**, 144106 (2019).

[27] H. Koch, H. J. A. Jensen, P. Jørgensen, T. Helgaker, G. E. Scuseria, and H. F. Schaefer, J. Chem. Phys. **92**, 4924 (1990).

[28] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic-Structure Theory* (John Wiley & Sons, Ltd., New York, 2014), Chap. 13, pp. 648–723.

[29] M. Wu, S. Chen, S. Camp, K. J. Schafer, and M. B. Gaarde, J. Phys. B: At., Mol., Opt. Phys. **49**, 62003 (2016).

[30] A. Iserles, *A First Course in the Numerical Analysis of Differential Equations*, 2nd ed., Cambridge Texts in Applied Mathematics (Cambridge University Press, Cambridge, UK, 2008).

[31] F. J. Harris, Proc. IEEE **66**, 51 (1978).

[32] S. D. Folkestad, E. F. Kjønstad, R. H. Myhre, J. H. Andersen, A. Balbi, S. Coriani, T. Giovannini, L. Goletto, T. S. Haugland, A. Hutcheson, I.-M. Høyvik, T. Moitra, A. C. Paul, M. Scavino, A. S. Skeidsvoll, Å. H. Tveten, and H. Koch, J. Chem. Phys. **152**, 184103 (2020).

[33] E. Tiesinga, P. J. Mohr, D. B. Newell, and B. N. Taylor, The 2018 CODATA recommended values of the fundamental physical constants, NIST standard reference database 121, web version 8.1, 2020.

[34] H. R. Reiss, Prog. Quantum Electron. **16**, 1 (1992).

[35] R. D. Johnson III, NIST computational chemistry comparison and benchmark database, NIST standard reference database 101, release 20, 2019.

[36] S. Coriani and H. Koch, J. Chem. Phys. **143**, 181103 (2015).

[37] T. H. Dunning, J. Chem. Phys. **90**, 1007 (1989).

[38] R. A. Kendall, T. H. Dunning, and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).

[39] E. Coccia, B. Mussard, M. Labeye, J. Caillat, R. Taïeb, J. Toulouse, and E. Luppi, Int. J. Quantum Chem. **116**, 1120 (2016).

[40] M. Labeye, F. Zapata, E. Coccia, V. Véniard, J. Toulouse, J. Caillat, R. Taïeb, and E. Luppi, J. Chem. Theory Comput. **14**, 5846 (2018).

[41] M. Bixon and J. Jortner, J. Chem. Phys. **48**, 715 (1968).

[42] T. B. Pedersen and H. Koch, J. Chem. Phys. **108**, 5194 (1998).

Paper III

Simulating weak-field attosecond processes with a Lanczos reduced basis approach to time-dependent equation-of-motion coupled-cluster theory

Andreas S. Skeidsvoll [1,*] Torsha Moitra [2,*] Alice Balbi [3] Alexander C. Paul [1]
Sonia Coriani [2,1,†] and Henrik Koch [1,3,‡]

[1]*Department of Chemistry, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway*
[2]*Department of Chemistry, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark*
[3]*Scuola Normale Superiore, Piazza dei Cavalieri 7, IT-56126 Pisa, Italy*

A time-dependent equation-of-motion coupled-cluster singles and doubles (TD-EOM-CCSD) method is implemented, which uses a reduced basis calculated with the asymmetric band Lanczos algorithm. The approach is used to study weak-field processes in small molecules induced by ultrashort valence pump and core probe pulses. We assess the reliability of the procedure by comparing TD-EOM-CCSD absorption spectra to spectra obtained from the time-dependent coupled-cluster singles and doubles method, and observe that spectral features can be reproduced for several molecules, at much lower computational times. We discuss how multiphoton absorption and symmetry can be handled in the method, and general features of the core-valence separation projection technique. We also model the transient absorption of an attosecond x-ray probe pulse by the glycine molecule.

I. INTRODUCTION

Stimulated by the recent experimental realization of various laser pulses with durations on the attosecond (1×10^{-18} s) time scale [1–5], capable of monitoring electronic motion, the theoretical simulation of coherent electron dynamics is currently an active field of research [6].

Real-time electronic structure theory considers the explicit time dependence of the electronic system by evolving the time-dependent Schrödinger equation in the time domain [6]. Explicitly time-dependent methods can directly provide the time-domain evolution of electronic wave functions together with nuclear motion, representing a versatile way of tracking ultrafast phenomena in both perturbative and nonperturbative regimes [7,8].

The development of real-time methods commenced in the late 1970s and early 1980s in the field of nuclear physics [9–11]. Despite these early endeavors, real-time methods did not become practical at that time due to the lack of electron correlation effects at the Hartree-Fock level and the high computational cost associated with propagation of correlated wave functions. However, decades of steady advancements in computing power and numerical algorithms have led to a renewed interest in explicit time propagation in correlated methods like density-functional theory [12,13], multiconfigurational self-consistent-field [14–16], configuration-interaction [17–20], algebraic diagrammatic construction [21,22], and coupled-cluster [23–33] methods.

In this paper, we present an implementation and representative case studies of the time-dependent equation-of-motion coupled-cluster (TD-EOM-CC) model for simulating weak-field attosecond valence pump–core probe processes. In conjunction with a reduced-space band Lanczos algorithm for obtaining the valence and core excited states, this model offers results similar to its time-dependent coupled-cluster (TDCC) counterpart in weak fields, at significantly lower computational costs. The reduction in cost enables the study of larger systems.

The paper is organized as follows. In Sec. II we detail the theory behind TD-EOM-CC and the asymmetric band Lanczos algorithm. Here, we also discuss a strategy used in order to guide the reduced space solver to directly obtain the transitions between excited states. The computational procedure used is detailed in Sec. III. In Sec. IV, simulations for various molecular systems are presented. First a benchmark study is presented for LiF, validating our proposed method. Second, the applicability of the core-valence separation (CVS) scheme is tested for LiH. Then, a two-photon absorption phenomenon has been captured using a stepwise procedure emulating the actual physical process for C_2H_4. Finally, we put forward a theoretical assessment of pump-probe absorption for the glycine molecule, which is deemed suitable for further experimental investigations. The findings are summarized in Sec. V.

II. THEORY

A. System

We model the system, composed of a molecule and its interaction with laser pulses, with the Hamiltonian

$$H(t) = H^{(0)} + V(t), \qquad (1)$$

---

*These authors contributed equally to this work.
†soco@kemi.dtu.dk
‡henrik.koch@sns.it

where the field-free Hamiltonian $H^{(0)}$ describes the molecule with fixed nuclei and without interactions with the external electromagnetic field. The semiclassical time-dependent interaction term, written in the dipole approximation and length gauge, is

$$V(t) = -\boldsymbol{d} \cdot \boldsymbol{\mathcal{E}}(t) \qquad (2)$$

and describes the interaction between the molecular electrons and the external electromagnetic field. The latter is represented by the electric field $\boldsymbol{\mathcal{E}}(t) = [\mathcal{E}_x(t)\, \mathcal{E}_y(t)\, \mathcal{E}_z(t)]^T$ and electronic dipole operator vectors, $\boldsymbol{d} = [d_x\, d_y\, d_z]^T$. We assume that the molecule is initially in the ground state of the field-free Hamiltonian, and take the electric field to be a linear combination of the electric fields of any number of laser pulses:

$$\boldsymbol{\mathcal{E}}(t) = \sum_n \boldsymbol{\mathcal{E}}_{0n} \cos[\omega_{0n}(t - t_{0n}) + \phi_n] f_n(t). \qquad (3)$$

The field of laser pulse n has an associated carrier frequency ω_{0n}, peak strength $|\boldsymbol{\mathcal{E}}_{0n}|$, polarization $\boldsymbol{\mathcal{E}}_{0n}/|\boldsymbol{\mathcal{E}}_{0n}|$, and an $8\sigma_n$-truncated Gaussian envelope function

$$f_n(t) = \begin{cases} e^{-(t-t_{0n})^2/(2\sigma_n^2)} & \text{if } |t - t_{0n}| \leqslant 8\sigma_n \\ 0 & \text{otherwise} \end{cases} \qquad (4)$$

with duration specified by σ_n, the temporal rms width. It is also specified by the central time t_{0n} and the carrier-envelope phase ϕ_n. We assume the carrier-envelope phase to be zero for all pulses, meaning that the maximum values of the envelope and cosine carrier functions belonging to pulse n coincide at t_{0n}.

The energy absorbed during the interaction can be given by [28,34]

$$\Delta E = \int_0^\infty \omega S(\omega) d\omega, \qquad (5)$$

where $S(\omega)$ is the response function

$$S(\omega) = -2\text{Im}[\langle\widetilde{\boldsymbol{d}}\rangle(\omega) \cdot \widetilde{\boldsymbol{\mathcal{E}}}^*(\omega)] \quad \omega > 0. \qquad (6)$$

The vectors $\langle\widetilde{\boldsymbol{d}}\rangle(\omega)$ and $\widetilde{\boldsymbol{\mathcal{E}}}(\omega)$ are the Fourier transforms of the time-dependent dipole moment expectation value and electric-field vectors, respectively, and the asterisk denotes complex conjugation. A positive or negative value of the function $S(\omega)$ describes the probability of absorption or emission of light with frequency ω, respectively [34].

B. TD-EOM-CC states

The time-dependent ket and bra of a TD-EOM-CC state can be expressed as

$$|\Psi(t)\rangle = \sum_j |\psi_j\rangle s_j(t), \quad \langle\widetilde{\Psi}(t)| = \sum_i k_i(t)\langle\widetilde{\psi}_i|, \quad (7)$$

where the italic indices i and j are used to denote general equation-of-motion coupled-cluster (EOM-CC) states, including the ground state with index zero. The time-independent EOM-CC kets and bras are given by

$$|\psi_j\rangle = e^T R_j |\text{HF}\rangle, \quad \langle\widetilde{\psi}_i| = \langle\text{HF}|L_i e^{-T}. \quad (8)$$

We assume that the EOM-CC states are biorthonormal:

$$\langle\widetilde{\psi}_i|\psi_j\rangle = \delta_{ij}. \qquad (9)$$

In the following, we let the indices κ and λ denote general determinants in the projection space, including the reference Hartree-Fock determinant with index zero. We use the indices μ and ν, on the other hand, to denote excited determinants.

The cluster operator T and the right and left operators R_j and L_i of Eq. (8) can be expressed as linear expansions in a finite set of operators τ_λ and $\tau_\kappa^\dagger$,

$$T = \sum_\nu \tau_\nu t_\nu, \quad R_j = \sum_\lambda \tau_\lambda r_{\lambda j}, \quad L_i = \sum_\kappa l_{i\kappa} \tau_\kappa^\dagger, \quad (10)$$

where the operator with index zero is the unit operator,

$$\tau_0 = \tau_0^\dagger = 1, \qquad (11)$$

and the τ_ν and $\tau_\mu^\dagger$ operators generate excited determinants from the ket and bra reference Hartree-Fock determinants, respectively:

$$\tau_\nu|\text{HF}\rangle = |\nu\rangle, \quad \langle\text{HF}|\tau_\mu^\dagger = \langle\mu|, \qquad (12)$$

$$\tau_\mu^\dagger|\text{HF}\rangle = 0, \qquad \langle\text{HF}|\tau_\nu = 0. \qquad (13)$$

We assume that the determinants are biorthogonal:

$$\langle\kappa|\lambda\rangle = \delta_{\kappa\lambda}. \qquad (14)$$

If all possible electronic excitations are included in the summations in Eq. (10), the method is equivalent to full configuration interaction. The sum can also be restricted to given excitation levels, giving approximate methods that scale polynomially with the system size. This includes the coupled-cluster singles and doubles method, where summation is only done over single and double excitations. We do not explicitly state the excitation levels included in the following expressions, since they hold for both restricted and unrestricted summation.

The cluster amplitudes t_ν in Eq. (10) can be found from solving equations involving the similarity-transformed field-free Hamiltonian operator $\bar{H}^{(0)}$ projected onto the right reference and left excited determinants

$$\langle\mu|\bar{H}^{(0)}|\text{HF}\rangle = 0, \qquad (15)$$

where the similarity transformation of an operator X is denoted by an overbar:

$$\bar{X} = e^{-T} X e^T. \qquad (16)$$

After the optimal cluster amplitudes t_ν have been determined, the right and left vectors of EOM-CC state i, with components $r_{\lambda j}$ and $l_{i\kappa}$, can be found as right and left eigenvectors of the field-free Hamiltonian matrix, with elements

$$H_{\kappa\lambda}^{(0)} = \langle\kappa|\bar{H}^{(0)}|\lambda\rangle. \qquad (17)$$

The right and left eigenvectors of the matrix in Eq. (17) with the lowest eigenvalue, specifying the ground EOM-CC state with index zero, have the following structure:

$$r_{00} = 1, \quad r_{\nu 0} = 0, \qquad (18)$$

$$l_{00} = 1, \quad l_{0\mu} = \bar{t}_\mu. \qquad (19)$$

The multipliers $\bar{t}_\mu$ are solutions to the equations

$$\langle\text{HF}|\bar{H}^{(0)}|\nu\rangle + \sum_\mu \bar{t}_\mu A_{\mu\nu}^{(0)} = 0, \qquad (20)$$

where elements of the field-free coupled-cluster Jacobian matrix $\boldsymbol{A}^{(0)}$ are given by

$$A_{\mu\nu}^{(0)} = \langle \mu | [\bar{H}^{(0)}, \tau_\nu] | \text{HF} \rangle. \tag{21}$$

The other right and left eigenvectors of the matrix in Eq. (17) correspond to excited EOM-CC states, denoted by the italic indices m and n. The eigenvectors have the following reference determinant components:

$$r_{0n} = -\sum_\nu \bar{t}_\nu r_{\nu n}, \tag{22}$$

$$l_{m0} = 0. \tag{23}$$

These components enforce the biorthogonality between the ground and excited states, in accordance with Eq. (9). The vectors $\boldsymbol{R}_m$ and $\boldsymbol{L}_m$, containing the components $r_{\nu m}$ and $l_{m\mu}$ of excited EOM-CC state m, are right and left eigenvectors of $\boldsymbol{A}^{(0)}$ with eigenvalue ω_m.

C. Derivation of TD-EOM-CC equations

The time derivative of the coefficients of the TD-EOM-CC ket can be found from projecting the ket time-dependent Schrödinger equation (TDSE)

$$\iota \frac{\partial}{\partial t} |\Psi(t)\rangle = H(t) |\Psi(t)\rangle, \tag{24}$$

where ι denotes the imaginary unit, onto the bra of EOM-CC state i, giving

$$\iota \frac{\partial s_i(t)}{\partial t} = \sum_j H_{ij}(t) s_j(t), \tag{25}$$

where the matrix elements of an operator $X(t)$ are given by

$$X_{ij}(t) = \langle \widetilde{\psi}_i | X(t) | \psi_j \rangle. \tag{26}$$

Likewise, the time derivative of the coefficients of the TD-EOM-CC bra can be found from projecting the bra TDSE,

$$-\iota \frac{\partial}{\partial t} \langle \widetilde{\Psi}(t) | = \langle \widetilde{\Psi}(t) | H(t), \tag{27}$$

onto the ket of EOM-CC state j, giving

$$-\iota \frac{\partial k_j(t)}{\partial t} = \sum_i k_i(t) H_{ij}(t). \tag{28}$$

The TD-EOM-CC equations (25) and (28) were to our knowledge first presented in Ref. [18], and have also been used in Ref. [19]. In those works, the matrix elements d_{ij} of the dipole moment operator, entering in the time-dependent Hamiltonian, are approximated by discarding non-Hermitian components. This was achieved by using the Hermitian $(d_{ij} + d_{ji}^*)/2$ instead of the d_{ij} given by Eq. (26). In the present paper, however, the full non-Hermitian d_{ij} are used in the solution of Eqs. (25) and (28). After the time-dependent coefficients $k_i(t)$ and $s_i(t)$ have been obtained, the time-dependent expectation value of a time-independent operator X can be calculated according to

$$\langle X \rangle(t) = \sum_{ij} k_i(t) X_{ij} s_j(t). \tag{29}$$

D. Asymmetric band Lanczos algorithm

We use the asymmetric band Lanczos algorithm to generate approximate eigenvalues $\widetilde{\omega}_n$ and right $\widetilde{\boldsymbol{R}}_n$ and left $\widetilde{\boldsymbol{L}}_n$ eigenvectors of the field-free Jacobian matrix $\boldsymbol{A}^{(0)}$. As outlined in Sec. II F, the approximate eigenvectors are used as a reduced basis for solving Eqs. (25) and (28).

The algorithm is a generalization of the simple asymmetric Lanczos algorithm, employing m right $(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_m)$ and p left $(\boldsymbol{c}_1, \ldots, \boldsymbol{c}_p)$ starting vectors instead of single ones [35–37]. A sequence of right vectors is constructed by transforming the right starting vectors by increasing powers of a given square asymmetric matrix $\boldsymbol{M}$. For our purpose, $\boldsymbol{M} = \boldsymbol{A}^{(0)}$. The i first vectors in the sequence, which can be linearly dependent, span the n-dimensional right band Krylov subspace,

$$\mathcal{K}_i(\boldsymbol{M}, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_m)$$
$$= \text{span}\{\underbrace{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_m, \boldsymbol{M}\boldsymbol{b}_1, \ldots, \boldsymbol{M}\boldsymbol{b}_m, \boldsymbol{M}^2\boldsymbol{b}_1, \ldots}_{i}\}, \tag{30}$$

where $i - n$ is the number of redundant vectors in the sequence. Likewise, a sequence of left vectors is constructed by transforming the left starting vectors by increasing powers of the transpose matrix $\boldsymbol{M}^T$. The j first vectors in the sequence, which can be linearly dependent, span the n-dimensional left band Krylov subspace,

$$\mathcal{K}_j(\boldsymbol{M}^T, \boldsymbol{c}_1, \ldots, \boldsymbol{c}_p)$$
$$= \text{span}\{\underbrace{\boldsymbol{c}_1, \ldots, \boldsymbol{c}_p, \boldsymbol{M}^T\boldsymbol{c}_1, \ldots, \boldsymbol{M}^T\boldsymbol{c}_p, (\boldsymbol{M}^T)^2\boldsymbol{c}_1, \ldots}_{j}\}, \tag{31}$$

where $j - n$ is the number of redundant vectors in the sequence. Note that the subspaces can be regarded as block Krylov subspaces [37] whenever i and j are multiples of the number of starting vectors. The n right $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ and left $(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n)$ Lanczos vectors, which form respective bases for $\mathcal{K}_i(\boldsymbol{M}, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_m)$ and $\mathcal{K}_j(\boldsymbol{M}^T, \boldsymbol{c}_1, \ldots, \boldsymbol{c}_p)$, are obtained by discarding redundant vectors of the sequences in Eqs. (30) and (31). A $n \times n$-dimensional band Krylov subspace approximation of $\boldsymbol{M}$ can then be obtained by expressing the matrix in the Lanczos vector bases.

The right and left Lanczos vectors can be generated iteratively with the recurrence relations [35]

$$\boldsymbol{M}\boldsymbol{V}_n = \boldsymbol{V}_n\boldsymbol{T}_n + \hat{\boldsymbol{V}}_n^C + \hat{\boldsymbol{V}}_n^D, \tag{32}$$

$$\boldsymbol{M}^T\boldsymbol{W}_n = \boldsymbol{W}_n\widetilde{\boldsymbol{T}}_n + \hat{\boldsymbol{W}}_n^C + \hat{\boldsymbol{W}}_n^D, \tag{33}$$

where the right and left Lanczos vectors form the matrices $\boldsymbol{V}_n = [\boldsymbol{v}_1 \quad \cdots \quad \boldsymbol{v}_n]$ and $\boldsymbol{W}_n = [\boldsymbol{w}_1 \quad \cdots \quad \boldsymbol{w}_n]$, respectively. The un-normalized vectors that form the nonzero columns of $\hat{\boldsymbol{V}}_n^C = [\boldsymbol{0} \quad \cdots \quad \boldsymbol{0} \quad \hat{\boldsymbol{v}}_{n+1} \quad \cdots \quad \hat{\boldsymbol{v}}_{n+m_c}]$ and $\hat{\boldsymbol{W}}_n^C = [\boldsymbol{0} \quad \cdots \quad \boldsymbol{0} \quad \hat{\boldsymbol{w}}_{n+1} \quad \cdots \quad \hat{\boldsymbol{w}}_{n+p_c}]$ serve as candidates for the next right and left Lanczos vectors, respectively. The sparse matrices $\hat{\boldsymbol{V}}_n^D$ and $\hat{\boldsymbol{W}}_n^D$ contain un-normalized candidates from previous iterations that have been deflated (i.e., discarded) due to linear dependence on already accepted right and left Lanczos vectors, respectively. Finally, the nonzero elements of $\boldsymbol{T}_n$ and $\widetilde{\boldsymbol{T}}_n$ are used to enforce the biorthogonality

between the $m + p + 1$ vectors that can overlap in exact arithmetic at each iteration [35].

In numerical implementations, vectors are usually deflated when linear independence is below a given threshold, since inexact arithmetic prevents the description of exact linear dependence. The numbers m_c and p_c, initially equal to the number of right and left starting vectors m and p, give the current number of vectors available for deflation. We say that the sequence in Eqs. (30) and (31) is fully exhausted when m or p deflations have occurred, respectively. The iterative procedure is then terminated, giving equal numbers of right and left Lanczos vectors.

The iterative solution of Eqs. (32) and (33) is done in accordance with Algorithm 5.1 of Ref. [37], with two exceptions. The following biorthogonalization step is added at the beginning of step 1:

if $n > 1$ **then**
    for $k = 1$ **to** $\max\{1, n - p_c - 1\}$ **do**
        $\hat{v}_n \leftarrow \hat{v}_n - v_k(w_k^T \hat{v}_n)$
    end for
end if

and the following at the beginning of step 2:

if $n > 1$ **then**
    for $k = 1$ **to** $\max\{1, n - m_c - 1\}$ **do**
        $\hat{w}_n \leftarrow \hat{w}_n - (\hat{w}_n^T v_k) w_k$
    end for
end if

These additions lead to an algorithm that enforces the biorthogonality between all Lanczos vectors in inexact arithmetic,

$$W_n^T V_n = \Delta_n = \mathrm{diag}(\delta_1, \delta_2, \ldots, \delta_m), \qquad (34)$$

and not just between the vectors that can overlap in exact arithmetic. We observe that this modification of the algorithm is important for numerical stability when the number of iterations becomes large, but the modification also makes the number of vector operations substantially higher. The number of operations can potentially be reduced in future implementations, e.g., by formulating a restarted asymmetric band Lanczos algorithm, based on existing approaches [38].

The iterative procedure continues until a given maximum chain length (i.e., number of iterations) $n = n^{\max}$ is reached, unless the procedure is terminated at a lower n because of full exhaustion of the sequence in Eq. (30) or (31). The algorithm generates the $n \times n$ matrices

$$T_n^P = \Delta_n^{-1} W_n^T M V_n$$
$$= T_n + \Delta_n^{-1} W_n^T V_n^D, \qquad (35)$$

$$\widetilde{T}_n^P = (W_n^T M V_n \Delta_n^{-1})^T$$
$$= \widetilde{T}_n + [(W_n^D)^T V_n \Delta_n^{-1}]^T, \qquad (36)$$

in accordance with Algorithm 5.1 of Ref. [37]. The matrices T_n^P and $\widetilde{T}_n^P$ are related by

$$\Delta_n T_n^P = (\widetilde{T}_n^P)^T \Delta_n, \qquad (37)$$

and are banded when no deflations have occurred [35].

The matrix T_n^P can be viewed as the oblique projection of M onto the n-dimensional $\mathcal{K}_i(M, b_1, \ldots, b_m)$ and orthogonally to the n-dimensional $\mathcal{K}_j(M^T, c_1, \ldots, c_p)$ [37]. Diagonalization of the matrix yields n eigenvalues, which approximate the eigenvalues of M, and associated right and left eigenvectors. The right eigenvectors can be transformed to approximate eigenvectors of M by premultiplication by V_n, and the left eigenvectors to approximate eigenvectors of M^T by premultiplication by $W_n^T \Delta_n^{-1}$ [35]. Approximate eigenvectors with dominant (low- and high-lying) eigenvalues are typically better converged than the ones in the middle [39–41].

E. Choice of starting vectors

In Appendix A, we demonstrate that operator matrix elements involving excited state n are linear in the right and left vector components $r_{\nu n}$ and $l_{n\mu}$. We assume that n is one of the states targeted by the band Lanczos algorithm, and state i is a previously calculated ground or excited state. Matrix elements between n and i can thus be written as the product of two vectors. The first vector is the right or left vector of state n. The second vector, which we take to be the starting vector of the algorithm, is based on state i and operator X. This choice of starting vectors simplifies the calculation of operator matrix elements, as will be shown in the following. Analogous arguments have been used in previous work [42–45] to guide the choice of start vectors for the simple Lanczos algorithm for coupled-cluster response [42,43] and EOM-CC theories [44,45].

The starting vectors based on the ground state $i = 0$, which are similar to the starting vectors used in coupled-cluster linear response theory [42,43], are given by

$$b_{\mu 0}^X = \xi_\mu^X, \qquad (38)$$

$$c_{0\nu}^X = {}^{\mathrm{EOM}}\eta_\nu^X - X_{00} \bar{t}_\nu, \qquad (39)$$

and the starting vectors based on excited EOM-CC state $i = m$ are given by [44]

$$b_{\mu m}^X = \sum_\nu \left({}^{\mathrm{EOM}}A_{\mu\nu}^X + \delta_{\mu\nu} X_{00} - \xi_\mu^X \bar{t}_\nu \right) r_{\nu m}, \qquad (40)$$

$$c_{m\nu}^X = \sum_\mu l_{m\mu} \left({}^{\mathrm{EOM}}A_{\mu\nu}^X + \delta_{\mu\nu} X_{00} - \xi_\mu^X \bar{t}_\nu \right). \qquad (41)$$

The specification of ξ_μ^X, ${}^{\mathrm{EOM}}\eta_\nu^X$, X_{00}, and ${}^{\mathrm{EOM}}A_{\mu\nu}^X$ is given in Appendix A.

The starting vectors can be expressed in the Lanczos basis by inserting the resolution of identity in terms of the biorthonormal Lanczos vectors:

$$b_i^X = \sum_j v_j \left(w_j^T b_i^X \right)$$
$$= \sum_{j=1}^m v_j b_{ji}^X, \qquad (42)$$

$$\left(c_i^X \right)^T = \sum_j \left[\left(c_i^X \right)^T v_j \right] w_j^T$$
$$= \sum_{j=1}^p c_{ij}^X w_j^T. \qquad (43)$$

The sum in Eq. (42) is restricted since $b_i^X \in \mathrm{span}\{v_1, \ldots, v_m\}$ while each $(w_{m+1}, \ldots)$ is biorthogonal to all $(v_1, \ldots, v_m)$. Likewise, the sum in Eq. (43) is restricted since $c_i^X \in \mathrm{span}\{w_1, \ldots, w_p\}$ while each $(v_{p+1}, \ldots)$ is biorthogonal to all $(w_1, \ldots, w_p)$.

Thus, the transition moments involving excited states can be obtained by contracting the starting vectors with the vectors of excited state n:

$$
\begin{aligned}
X_{in} &= \left(c_i^X\right)^T R_n \\
&= \sum_{j=1}^p c_{ij}^X \left(w_j^T R_n\right) \\
&= \sum_{j=1}^p c_{ij}^X R_{jn}, \quad\quad\quad (44)
\end{aligned}
$$

$$
\begin{aligned}
X_{ni} &= L_n^T b_i^X \\
&= \sum_{j=1}^m \left(L_n^T v_j\right) b_{ji}^X \\
&= \sum_{j=1}^m L_{nj} b_{ji}^X. \quad\quad\quad (45)
\end{aligned}
$$

R_{jn} and L_{nj} are simply the components of the right and left eigenvectors of T_n^P, respectively, and $b_{ji}^X = w_j^T b_i^X$ and $c_{ij}^X = (c_i^X)^T v_j$ are products of starting vectors and biorthonormal Lanczos vectors.

F. Generation of a reduced basis

The iterative process that is used to calculate sets of excited EOM-CC states is given below. The set J_c contains the indices of already calculated EOM-CC states. At the beginning of the procedure, where $c = 0$, only the ground state has been calculated, and $J_0 = \{0\}$. The iterative procedure for the cth EOM-CC state calculation is as follows.

(1) Choose a subset of the state indices from previous calculations, $I_c \subseteq J_{c-1}$, and a set of operators X_c based on the final states that can be accessed by the operators (see Sec. IV). Also choose a maximum chain length n_c^{max}, and maximum eigenvalue ω_c^{max} and minimum transition strength S_c^{min} values.

(2) Sequences of right and left starting vectors, $(b_i^X)_{i \in I_c, X \in X_c}$ and $(c_i^X)_{i \in I_c, X \in X_c}$, are constructed in accordance with Eqs. (38)–(41).

(3) The band Lanczos algorithm described in Sec. II D is run with the maximum chain length n_c^{max}, the field-free Jacobian matrix $A^{(0)}$, and the sequences of starting vectors. The algorithm terminates at $n \leqslant n_c^{\mathrm{max}}$, constructing the matrix $T_{n_c}^P$.

(4) The eigenvalues and corresponding right and left eigenvectors of $T_{n_c}^P$ are calculated. Together, these determine a set of n_c approximate EOM-CC states indexed by N_c.

(5) States $n \in N_c$ with approximate eigenvalues $\tilde{\omega}_n$ of $A^{(0)}$ that are greater than ω_c^{max} are discarded.

(6) Matrix elements for all operators $X \in X_c$ and combinations of final $n \in N_c$ and initial $i \in I_c$ states are calculated in accordance with Eqs. (44) and (45). States $n \in N_c$ with

transition strengths $S_{in}^X = X_{in} X_{ni}$ that are smaller than S_c^{min} for all operators and initial states are discarded.

(7) For each nondiscarded state $n \in N_c$, the right and left eigenvectors of $T_{n_c}^P$ are transformed to approximate right and left eigenvectors $\tilde{R}_n$ and $\tilde{L}_n$ of $A^{(0)}$ by premultiplication by V_n and $W_n^T \Delta_n^{-1}$, respectively.

(8) If an assessment of the convergence of the vectors is requested, residual norms of all approximate right and left vectors are calculated. All states with a left or right residual norm exceeding a given threshold are discarded.

(9) Finally, nondiscarded states with vectors linearly independent of previously calculated vectors are stored, and indexed by $N_c' \subseteq N_c$. The indices are added to the previous index set $J_c = J_{c-1} \cup N_c'$.

The iterative procedure is repeated if states of higher excitation levels are desired. Afterwards, the excited-state Jacobian and overlap matrices, with elements $\tilde{A}_{mn}^{(0)} = \tilde{L}_m^T A^{(0)} \tilde{R}_n$ and $\tilde{S}_{mn} = \tilde{L}_m^T \tilde{R}_n$, are constructed in the reduced basis of the approximate eigenvectors. The right and left generalized eigenvalue problems are solved, giving new sets of right and left eigenvectors of $\tilde{A}^{(0)}$. The dipole and field-free Hamiltonian matrices are then calculated in the basis of both the ground and the newly generated excited states, in accordance with Eq. (26), and used in solving the time-dependent problems defined by Eqs. (25) and (28).

III. COMPUTATIONAL DETAILS

Experimental geometries from the NIST database [46] are used for LiH, LiF, and C_2H_4. An optimized geometry from the same database is used for glycine, obtained with the MP2 method with all electrons correlated and the cc-pVTZ basis set. The linear molecules LiH and LiF are aligned along the z axis, as done in Ref. [28]. The ethylene molecule is placed in the xy plane, with the C-C bond along the x axis. The glycine molecule is of C_s symmetry for the chosen geometry, with the xy plane as the mirror plane. An illustration of the structures of the molecules, together with the polarizations of the valence-exciting pump and core-exciting probe pulses, is shown in Fig. 1.

In all following calculations, the aug-cc-pCVDZ basis set [47,48] is used for atoms targeted by the core-exciting pulses; the aug-cc-pVDZ [47] basis set is adopted for the remaining
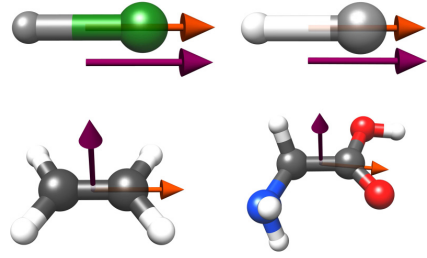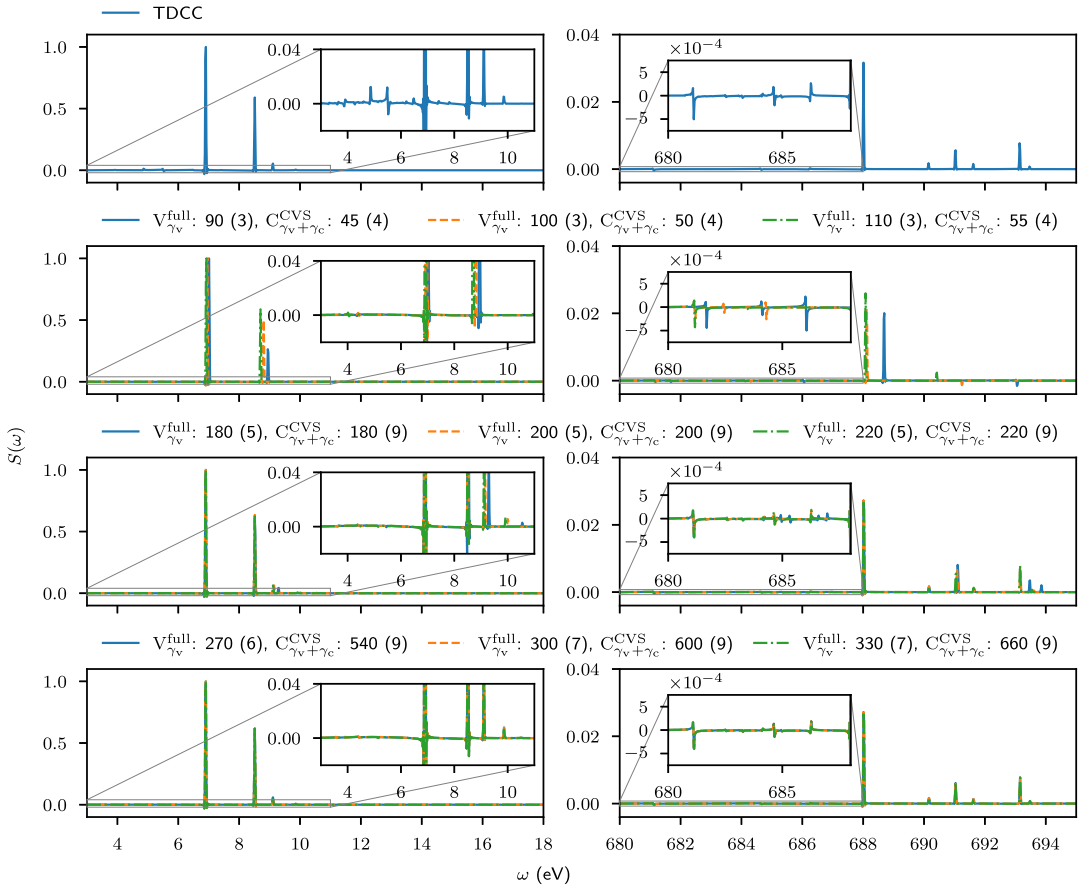

FIG. 1. Illustration of the structures of lithium fluoride (top left), lithium hydride (top right), ethylene (bottom left), and glycine (bottom right), together with the polarization of the valence-exciting pump (orange arrow) and core-exciting probe (purple arrow) pulses.

FIG. 2. LiF pump-probe absorption $S(\omega)$ as a function of frequency ω in the valence and core regions, normalized by the tallest peaks in the spectra. The TDCC results are shown in the top left and right panels. TD-EOM-CC results, calculated at different band Lanczos chain lengths, are shown in the lower panels. EOM-CC valence (V) states are calculated in the full projection space, while the core (C) states are calculated within the CVS approximation. Valence states energetically inaccessible by a single pump photon γ_v are discarded, and so are core states energetically inaccessible by subsequent absorption of a probe photon $\gamma_v + \gamma_c$. The chain lengths of the calculations are given, together with the number of converged states (in brackets).

atoms in the molecules. Valence and core states are obtained with the asymmetric band Lanczos algorithm with varying chain lengths as specified in Sec. IV.

Lanczos vectors with Euclidian norms of less than 1×10^{-9} are deflated, but this did not occur in any of the calculations. Final excited states that do not have a minimum transition strength of at least 1×10^{-7} to any initial state, for any of the operators used to construct the starting vectors, are discarded. This is done to only keep states that give a non-negligible contribution to the dynamics. Also, states with excitation energies above $\omega^{\max} = \sum_i^{n_\gamma} E_{\gamma_i}^{\max}$ are discarded, where $E_{\gamma_i}^{\max} = \omega_i + 8\sigma_i^{\omega}$ is an estimate of the maximum energy of photon i involved in the n_γ-photon transition to the desired excited states. The carrier frequency ω_i and the frequency rms width $\sigma_i^{\omega} = 1/(2\sigma_i^t)$ are parameters of the pulse providing photon i (see Sec. II A).

CVS [49–51] projectors were used to calculate core states. A "core-only" CVS projector is applied to remove excitations that originate exclusively from valence orbitals. This is done by zeroing out all right and left vector elements that only involve molecular orbitals with energies greater than the energy of the lowest core molecular orbital of a given atom. This yields a Lanczos spectrum starting at the lowest core excitation energy of the chosen edge. A complementary "valence-only" CVS projector is used to obtain valence excited states that are orthogonal to the core excited states.

Except for the spectra presented in Fig. 2, only sufficiently converged valence and core band Lanczos vectors are used for calculating stationary states and corresponding Hamiltonian and transition moment matrices. This is done by discarding states with either right or left residual norms greater than 1×10^{-2} for valence states and 1×10^{-1} for core states.

In all calculations, valence states are calculated first, with starting vectors based on the ground state. All accepted valence states are, together with the ground state, used to construct starting vectors for the core state calculations (see Sec. II F). The energies and maximum transition strengths for all accepted valence and core states are given in the Supplemental Material [52].

A fixed pump-probe delay of 40 a.u. (about 0.968 fs) is used for lithium fluoride, lithium hydride, and ethylene. The delay is varied for glycine, in order to calculate the transient absorption of the molecule. In all calculations the central time of the probe pulse is set to zero and the negative central times of the pump pulses are set accordingly.

Unless otherwise stated, integration of the TD-EOM-CC and TDCC equations is done using a Dormand-Prince 5(4) integration scheme [53] with a maximum time step of 0.1 a.u., and maximum and minimum local errors of 1×10^{-7} and 1×10^{-9}, respectively (see Appendix B). Each component of the time-dependent dipole moment expectation value and electric-field vectors is multiplied with the Hann window before Fourier transformation.

All calculations are performed using a development version of the e^T program [54].

IV. RESULTS AND DISCUSSION

A. Lithium fluoride: Convergence and nonlinear pump interaction

When discussing the applicability of the band Lanczos algorithm for modeling attosecond pump-probe processes, a key question is how spectra are affected by the chain length used. With this in mind, a single TDCC LiF pump-probe absorption spectrum, calculated with the RK4 integrator and fixed time steps of 5×10^{-3} a.u., is in Fig. 2 compared to TD-EOM-CC spectra calculated with the Dormand-Prince 5(4) integration scheme and various band Lanczos chain lengths. In all calculations, the pulses have the parameters used for the LiF spectra in Sec. III B of Ref. [28], where the F K edge is targeted by the probe pulse. Figure 1 illustrates the polarization of the pulses relative to the orientation of the molecule. All states with energies inaccessible by the absorption of one photon from each pulse are discarded from the TD-EOM-CC calculations.

For lower chain lengths, the peaks of the band Lanczos spectra shown in Fig. 2 both shift and scale significantly with variations in the chain length, indicating that excitation energies and dipole matrix elements are badly converged. The convergence generally improves with the chain length, and low-energy high-amplitude peaks seem to converge first. Higher chain lengths are needed for good convergence of high-energy low-amplitude peaks, as expected from the convergence behavior of Lanczos algorithms.

As demonstrated, the inclusion of badly converged states can give spectral peaks with incorrect positions and amplitudes. In addition, these states can also increase the cost of matrix element calculation and propagation, decrease the convergence rate of consecutive band Lanczos calculations, and cause serious numerical instabilities during propagation. In order to avoid these adverse effects, states with badly con-

verged right or left vector residual norms will be discarded in the following band Lanczos calculations.

In Fig. 3, the aforementioned TDCC LiF pump-probe absorption spectrum is compared to TD-EOM-CC spectra from converged states only. Note that the three most dominant peaks in the TDCC spectrum are present in the green spectrum, which is calculated with a valence chain length of 300, but a chain length of 400 is needed in order to converge the short peak at around 9.9 eV. The low amplitude peaks below and around the tall peak at around 6.9 eV are missing.

In an earlier work [28], we speculated that the smaller peaks below 6.9 eV in the pump-only LiF spectrum could originate from two-photon absorption. This claim was later discussed by Pedersen *et al.* [32], where the TDCC state of LiF interacting with the pump pulse was analyzed in terms of stationary state populations. Their analysis supports the interpretation that two photons are absorbed from the pump pulse.

In order to take two-photon absorption into account, spectra are recalculated with the inclusion of valence states energetically accessible by two pump photons and core states accessible by an additional probe photon. The corresponding results obtained with chain lengths of 300 and 400 are shown in purple and red in Fig. 3, respectively. Note that the 300 valence chain length spectrum still lacks the smaller features of the TDCC spectrum, but the 400 valence chain length spectrum is practically indistinguishable from the TDCC one. This similarity corroborates the claim that two photons are absorbed from the pump pulse. Furthermore, the results demonstrate that reduced-basis TD-EOM-CC can faithfully reproduce TDCC results in particular systems, even when nonlinear interactions are involved. The embedded Dormand-Prince 5(4) integrator is seen to perform well for TD-EOM-CC.

The bottom panel of Fig. 3 demonstrates the use of the valence-only CVS projector to calculate the valence states. The approximation seems to improve the rate of convergence with respect to chain length, as a length of 300 is enough to retrieve all the features of the TDCC spectrum while a higher number is necessary in the nonprojected case. This improved convergence can be explained by the reduction in dimension from projecting out transitions from core orbitals. Moreover, since the approximation does not seem to lead to significant scaling or shifting of the valence peaks, it is adopted in the following calculations.

B. Lithium hydride: Applicability of the CVS projectors

To further assess the performance of the proposed procedure, as well as the applicability of the CVS projectors, we use the TD-EOM-CC procedure to model the interaction of the lithium hydride molecule with the pump-probe pair described in Sec. III A of Ref. [28]. Figure 1 illustrates the polarization of the pulses relative to the orientation of the molecule. Li K-edge spectra are notoriously difficult to describe accurately due to the small energy separation between the valence and core excitation regions. This can be considered a challenging test case for the applicability of the core-valence separation scheme.

A comparison between TD-EOM-CC and TDCC spectra is given in Fig. 4, where the latter is taken from Ref. [28]. In all core state calculations a fixed band Lanczos chain length of
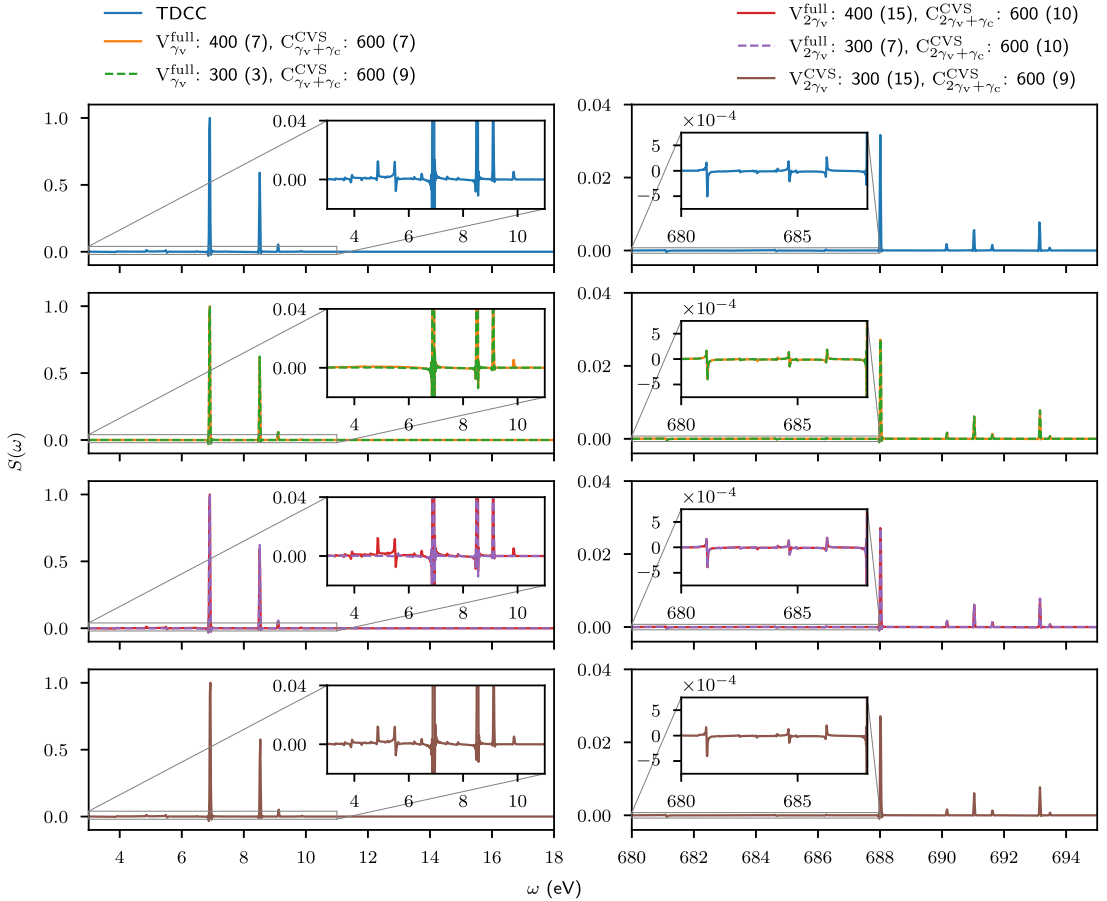
FIG. 3. LiF pump-probe absorption $S(\omega)$ as a function of frequency ω in the valence and core regions, normalized by the tallest peaks in the spectra. The TDCC results are shown in the top left and right panels. TD-EOM-CC results, calculated at different band Lanczos chain lengths, are shown in the lower panels. EOM-CC valence (V) states are calculated in the full projection space (middle panels) or within the CVS approximation (bottom panels), while the core (C) states are calculated within the CVS approximation only. Valence states energetically inaccessible by a single or two pump photons, γ_v or $2\gamma_v$, are discarded, and so are core states energetically inaccessible by subsequent absorption of a probe photon, $\gamma_v + \gamma_c$ or $2\gamma_v + \gamma_c$. The chain lengths of the calculations are given, together with the number of converged states (in brackets).

400 is used. However, the number of converged core states, given in brackets, differs due to the different starting vectors employed.

Since TD-EOM-CC with energy-limited valence and core states successfully reproduced the TDCC spectrum of LiF in Sec. IV A, a similar procedure is attempted for calculating the TD-EOM-CC LiH spectrum. That is, valence states inaccessible by two pump photons and core states inaccessible by an additional probe photon are discarded. The results are shown in orange in the second topmost panels of Fig. 4. The spectrum lacks some of the weaker features in the valence excitation energy region, and, more notably, many of the dominant features in the core excitation region. In other words, a characteristic of the LiH molecule seemingly prevents us from reproducing the TDCC spectrum using the procedure in

the previous section. In the following, we argue that the Li K edge in LiH involves states that cannot be obtained with the core-only CVS projector alone, since they do not correspond to core excitations.

TD-EOM-CC spectra calculated with states obtained with the valence-only CVS projector, energetically accessible by two pump and one probe photons, are shown in the three middle rows of panels in Fig. 4. The valence chain lengths used are 100 (green), 150 (red), and 200 (purple). The number of converged states in the valence region increases with the chain length. Remarkably, increasing the valence chain length also leads to additional peaks in the core region, illustrated in the right panels. This demonstrates that, apart from the two intense peaks obtained at about 54.1 eV and 57.7 eV, the other peaks are of pure valence excitation character.
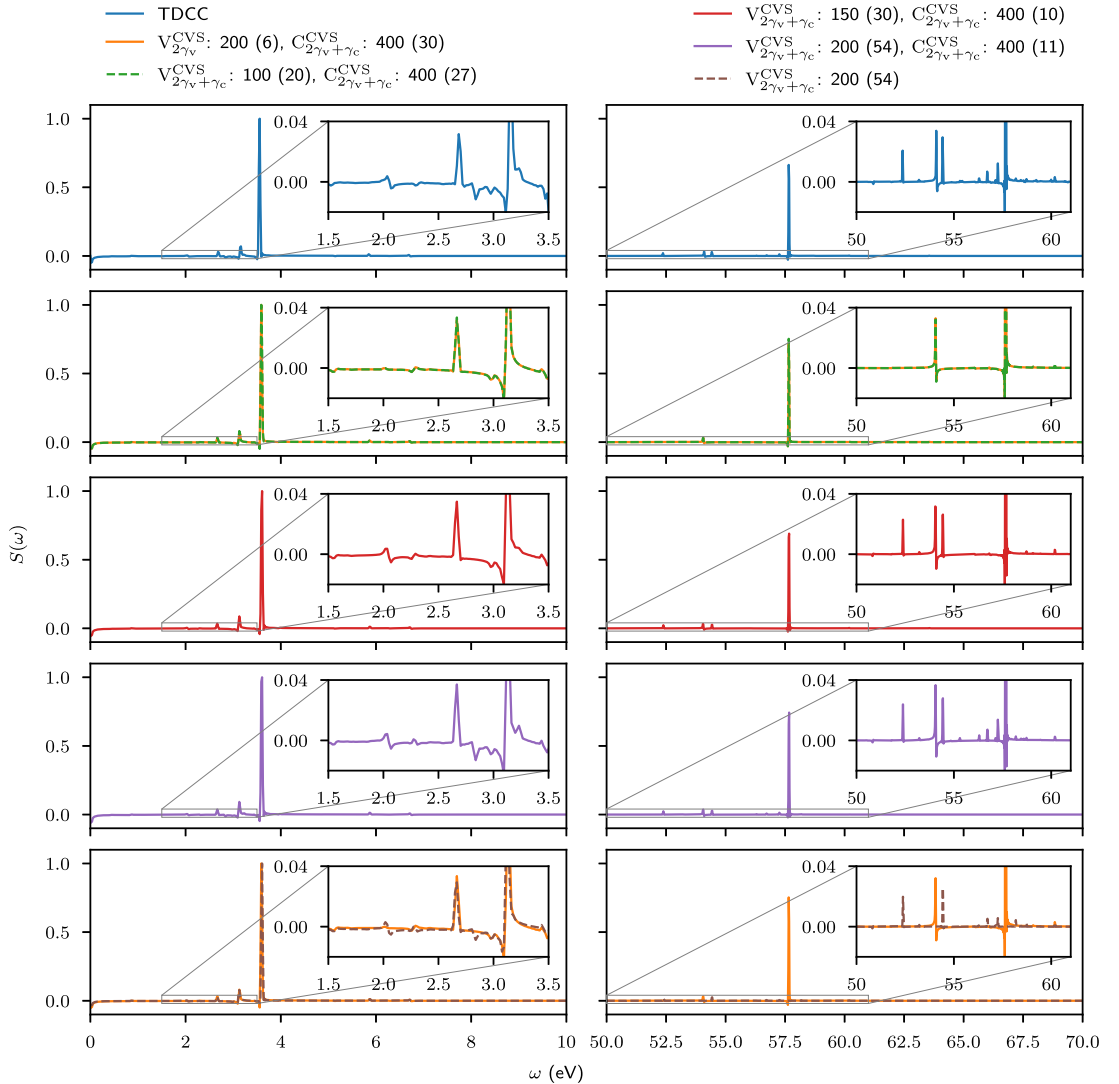
FIG. 4. LiH pump-probe absorption $S(\omega)$ as a function of frequency ω in the valence and core regions, normalized by the tallest peaks in the spectra. The TDCC results are shown in the top left and right panels. TD-EOM-CC results, calculated at different band Lanczos chain lengths, are shown in the lower panels. All EOM-CC valence (V) and core (C) states are calculated within the CVS approximation. For the CVS valence calculation for the results shown in orange, states energetically inaccessible by two pump photons, $2\gamma_v$ are discarded. For all other band Lanczos calculations, only the states inaccessible by two pump photons and one probe photon $2\gamma_v + \gamma_c$ are discarded. The results shown in brown are calculated from CVS valence states only. The chain lengths of the calculations are given, together with the number of converged states (in brackets).

The necessity to include high-energy states calculated with the valence-only CVS projector is further validated by superimposing a spectrum exclusively from valence-only CVS states (brown) with the spectrum calculated from energy-limited valence and core states (orange), shown in the bottom panels of Fig. 4. The peaks of the composite spectrum are in good agreement with the TDCC ones. Therefore, we assert that use of both the core-only and the complementary valence-only CVS projectors is necessary in order to accurately capture the spectral features around the Li K edge in LiH. Note that this should not be taken as a failure of the CVS projectors, but as a consequence of the peculiar electronic structure of LiH. In fact, the high-energy states of pure valence character can be more easily calculated in the dimension reduced by the valence-only CVS projector. One may still question whether the corresponding peaks will be as
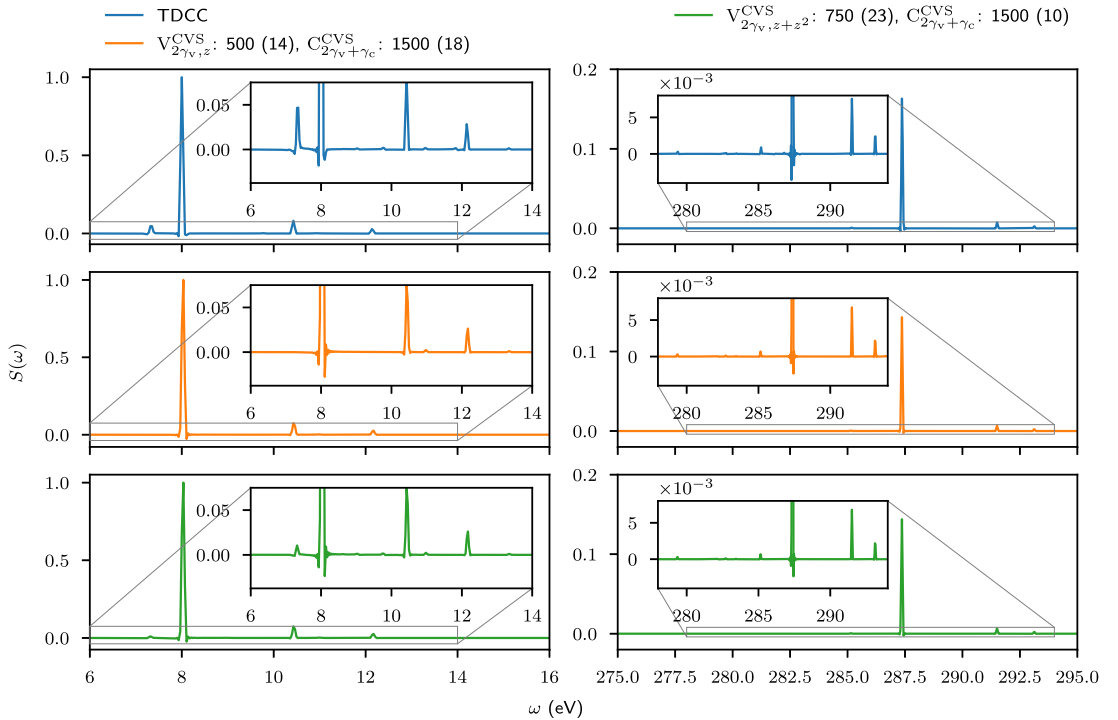
FIG. 5. Ethylene pump-probe absorption $S(\omega)$ as a function of frequency ω in the valence and core regions, normalized by the tallest peaks in the spectra. The TDCC results are shown in the top left and right panels, and TD-EOM-CC results are shown in the lower panels. In the middle panels, a ground-state z dipole operator starting vector is used for calculating the valence (V) states. In the bottom panels, both ground-state z dipole operator and z^2 quadrupole operator starting vectors are used. All valence and core (C) states are calculated within the CVS approximation. The chain lengths of the calculations are given, together with the number of converged states (in brackets).

prominent in experimental spectra, since continuum electrons are very crudely represented in the chosen basis set.

C. Ethylene: Nonlinear pump interaction for a different symmetry group

For ethylene, rms widths of the z-polarized pump and x-polarized probe pulses are set to 10 a.u. and 5 a.u., corresponding to intensity full width at half maximum (FWHM) durations of about 403 as and 201 as, respectively. Figure 1 illustrates the polarization of the pulses relative to the orientation of the molecule. The carrier frequency of the pump pulse is set to 0.294 114 89 a.u. (about 8.0 eV), and the probe pulse is set to 10.495 830 66 a.u. (about 285.6 eV, C K edge). The field strengths of the pump and probe pulses are set to 0.01 a.u. and 0.1 a.u., respectively. The time-dependent state is propagated with the Dormand Prince 5(4) integration scheme, from -2500 a.u. to 2500 a.u. of time. The TDCC spectrum, shown in the top panel Fig. 5, is characterized by four dominant peaks in the valence excitation region. A low amplitude peak at around 7.3 eV is present in the TDCC spectrum, but missing in the TD-EOM-CC spectrum calculated with a z dipole operator starting vector, shown in the middle panels. In accordance with the interpretation of the spectrum of LiF in Sec. IV B, we attribute the missing peak to a two-photon excitation process,

even though valence states energetically accessible by two photons are included. Note that quadratic functions of the z dipole operator belong to the A_g representation of D_{2h}, the point group of ethylene for the chosen geometry. Hence, we should not expect the single starting vector, belonging to the B_{1u} representation, to facilitate the convergence of the two-photon peaks.

In order to mimic the two pump photon absorption process, we include a starting vector constructed from the z^2 quadrupole operator in the valence-state calculation. The results, shown in the bottom panels of Fig. 5, now capture the two-photon peak at around 7.3 eV. The amplitude yielded by TD-EOM-CC is, however, underestimated compared to the TDCC one, which might indicate that more secondary valence excited states should be included in the computation. It might also hint at differences in two-photon absorption as described by TD-EOM-CC and TDCC.

D. Glycine: Transient absorption

As a final example, we use the computational procedure to model attosecond transient absorption by the glycine molecule. The polarization of the pump pulse is set to the polarization of the EOM-CC transition dipole moment between the ground and first dipole allowed valence excited
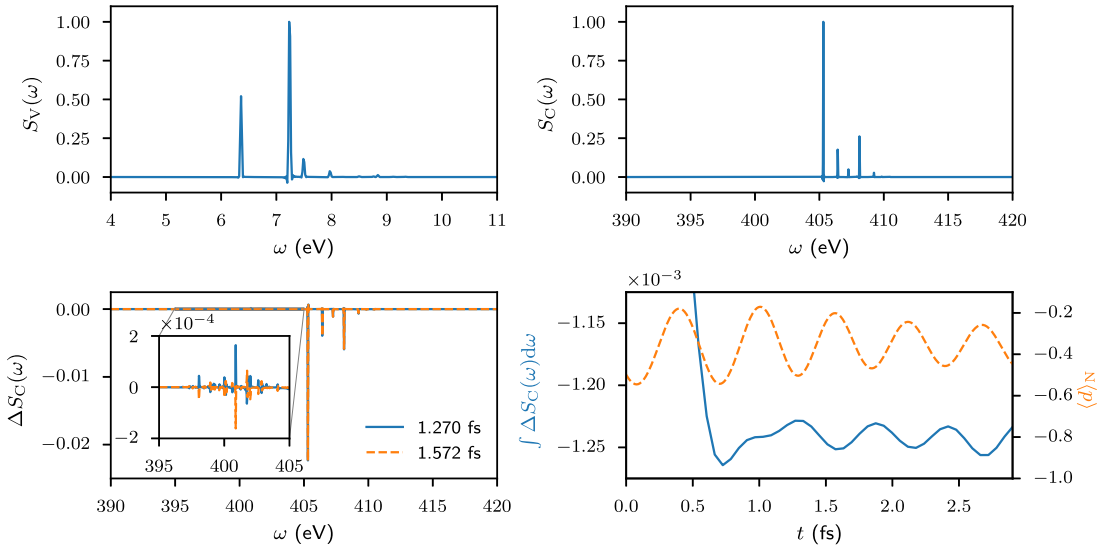
FIG. 6. Glycine. Top left: Pump pulse absorption from the ground state $S_V(\omega)$, normalized by the tallest peak in the spectrum. Top right: Probe pulse absorption from the ground state $S_C(\omega)$, normalized by the tallest peak in the spectrum. Bottom left: Pump-probe absorption minus the probe absorption from the ground state $\Delta S_C(\omega) = S(\omega) - S_C(\omega)$, normalized by the tallest peak in the $S_C(\omega)$ spectrum. The results for two different pump-probe delays are shown. Bottom right: In blue, the numerically integrated probe absorption difference $\int \Delta S_C(\omega)d\omega$ is shown as a function of pump-probe delay. In orange, the dipole induced by the pump pulse in the direction from the center of mass to the N atom, $\langle d \rangle_N$, is shown as a function of time after the center of the pump pulse.

state, $(0.490\,072x + 0.871\,682y)$, which is in the mirror plane of the molecule (xy plane). The probe is z polarized. Figure 1 illustrates the polarization of the pulses relative to the orientation of the molecule. The rms widths of the pump and probe pulses are set to 20 a.u. and 10 a.u., corresponding to intensity FWHM durations of about 806 as and 403 as, respectively. The carrier frequency of the pump pulse is set to 0.233 683 25 a.u. (about 6.4 eV), and the probe pulse is set to 14.894 573 19 a.u. (about 405.3 eV, N K edge). The field strengths of the pump and probe pulses are set to 0.01 a.u. and 0.1 a.u., respectively. The time-dependent state is propagated with the Dormand Prince 5(4) integration scheme, from −5000 a.u. to 5000 a.u. of time.

A single band Lanczos calculation is used for constructing the valence states, where states energetically inaccessible by two-photon transitions are discarded. Note that we do not need to use quadrupole operators in order to get two-photon valence states of glycine in the reduced basis, as otherwise done for the ethylene valence states, since both linear and quadratic functions of x and y dipole operators belong to the A' representation of C_s. A valence-state calculation, with ground-state starting vectors and a chain length of 1500, gives 17 converged states. A subsequent core state calculation, with ground- and valence-state starting vectors and a chain length of 3000, gives 20 converged states.

As a note of caution, all the converged valence states have energies below 10.5 eV, which is below the double frequency of the carrier photons. This indicates that two-photon absorption is not properly accounted for by the reduced basis, and

nonlinear features involving higher-energy valence states may be missing in the spectra.

In the top left panel of Fig. 6, the absorption of the pump pulse is shown as a function of frequency. Even though the glycine molecule is substantially larger than the other molecules considered, the spectrum is still dominated by a small number of peaks. The number of dominant peaks is also smaller than the number of converged states in the basis (=17). The spectrum of the absorption of the probe pulse by the ground state, shown in the top right panel, also has fewer dominant peaks than the number of converged states (=20).

In order to calculate the transient absorption of the probe pulse by the glycine molecule, absorption spectra are calculated with the pump-probe setup used for the other molecules, with pump-probe delays varying from 0 a.u. to 120 a.u. (about 2.903 fs) in intervals of 2.5 a.u. The reduced basis energies and dipole matrix elements do not have to be recalculated between the different TD-EOM-CC calculations, since these are independent of the pump-probe delay. Difference spectra are then calculated by subtracting the ground-state probe absorption spectrum from the pump-probe absorption spectra, before normalizing by the tallest peak in the ground-state probe spectrum. In the bottom left panel, the difference spectra at 52.5 a.u. and 65 a.u. (about 1.270 fs and 1.572 fs) are shown in blue and orange, respectively. Both spectra are dominated by negative peaks, indicative of ground-state bleaching. In addition, the spectra vary slightly with pump-probe delay, which is particularly visible for the peaks that are not energetically accessible from the ground state, e.g., in the energy range from 395 eV to 405 eV shown in the inset.

In order to quantify the delay-dependent difference in the absorption of the full probe pulse, each of the 49 pump-probe difference spectra are numerically integrated from 390 eV to 420 eV using the trapezoidal rule. The results are shown as function of pump-probe delay in the bottom right panel of Fig. 6, in blue. Note that the absorption difference is smaller for short pump-probe delays, which can be explained by the fact that ground-state bleaching happens gradually during the pump pulse interaction. A shorter pump-probe delay implies that the molecule is probed while bleaching still occurs, which can lead to a smaller difference between the pump-probe absorption and the ground-state probe absorption.

We have also calculated the pump-induced time-dependent dipole moment in the direction from the center of mass to the center of the N atom, as a way of quantifying the migration of charge between the end containing the N atom and the opposite end of the molecule. The dipole moment is shown in the bottom right panel of Fig. 6, in orange.

Note that the dominant periods of both the time-dependent dipole moment and the integrated absorption (after 1 fs), shown in the bottom right panel of Fig. 6, fall within 0.57(4) fs. This indicates that the pump-induced TD-EOM-CC state is a coherent superposition dominated by states with energy differences of 7.3(6) eV, which is in agreement with the ground-state pump absorption spectrum (top left). It also indicates that the dominant features of the time-dependent charge migration and the delay-dependent K-edge absorption are correlated and can be measured with phase-controlled pulses with finite duration, as has previously been demonstrated for instantaneous pulses [55,56].

V. CONCLUSION

We have demonstrated the use of the asymmetric band Lanczos algorithm to generate reduced TD-EOM-CC bases for various molecules, taking the characteristics of pulses suitable for probing attosecond phenomena into account. The specific starting vectors used in the calculations direct the band Lanczos algorithm towards states that are useful for representing the interactions. The starting vectors also allow for the affordable calculation of transition strengths, which are used, together with excitation energies, to automatically select the reduced basis. The basis is further reduced by removing unconverged states.

In Sec. IV A, we demonstrated how lithium fluoride spectral peaks can converge towards peaks calculated with TDCC by increasing the band Lanczos chain length and taking a sufficient number of relevant states into account. In particular, we showed that two-photon absorption has to be taken into account in order to reproduce the smaller features of the TDCC spectrum, as speculated in Ref. [28].

In Sec. IV B, we demonstrated that the core-only CVS projector eliminates several of the peaks around the K edge of lithium in lithium hydride. The missing peaks can be captured with the complementary valence-only CVS projector, which enabled us to target high-energy states of pure valence character. This observation indicates that care should be taken when the CVS scheme is used for light elements such as lithium, where the energy separation of the core and valence orbitals is small, so that pure valence excitations can fall within the region of core excitations.

In Sec. IV C, we used starting vectors constructed from both dipole and quadrupole operators, in order to converge ethylene valence states that are dark with respect to one-photon transitions from the ground state.

In Sec. IV D, pump and probe pulses with varying time delays were used to assess the transient absorption of a K-edge probe pulse as a function of pump-probe delay. We showed how the transient absorption seems to correlate with the migration of charge induced by the pump, and how both quantities seem to reveal the dominant timescale in the coherent superposition.

APPENDIX A: REWRITING EOM-CC MATRIX ELEMENT EXPRESSIONS

The matrix element X_{ij} of the operator X and the left and right vectors of EOM-CC states i and j, respectively, can be written as

$$
\begin{aligned}
X_{ij} &= \langle \widetilde{\psi}_i | X | \psi_j \rangle \\
&= \sum_{\kappa\lambda} l_{i\kappa} \langle \kappa | \bar{X} | \lambda \rangle r_{\lambda j} \\
&= \sum_{\nu} l_{i0} \langle \mathrm{HF} | \bar{X} | \nu \rangle r_{\nu j} + \sum_{\mu\nu} l_{i\mu} \langle \mu | \bar{X} | \nu \rangle r_{\nu j} \\
&\quad + \left(l_{i0} X_{00} + \sum_{\mu} l_{i\mu} \xi_\mu^X \right) r_{0j} \\
&= \sum_{\nu} l_{i0} \langle \mathrm{HF} | [\bar{X}, \tau_\nu] | \mathrm{HF} \rangle r_{\nu j}
\end{aligned}
$$

$$+ \sum_{\mu\nu} l_{i\mu} \left({}^{\mathrm{LR}}A_{\mu\nu}^X + \langle \mu | \tau_\nu \bar{X} | \mathrm{HF} \rangle \right) r_{\nu j}$$

$$+ \left(l_{i0} X_{00} + \sum_\mu l_{i\mu} \xi_\mu^X \right) r_{0j}, \tag{A1}$$

where

$$ {}^{\mathrm{LR}}A_{\mu\nu}^X = \langle \mu | [\bar{X}, \tau_\nu] | \mathrm{HF} \rangle, \tag{A2}$$

$$ X_{00} = \langle \mathrm{HF} | \bar{X} | \mathrm{HF} \rangle, \tag{A3}$$

$$ \xi_\mu^X = \langle \mu | \bar{X} | \mathrm{HF} \rangle. \tag{A4}$$

1. Excited-state left and ground-state right vectors

The left vector of excited state m has the reference component $l_{m0} = 0$ [see Eq. (23)], while the right vector of the ground state has components $r_{00} = 1$ and $r_{\nu 0} = 0$ [see Eq. (18)]. Inserting this into Eq. (A1), we obtain

$$ X_{m0} = \sum_\mu l_{m\mu} \xi_\mu^X, \tag{A5}$$

which is also the expression appearing in CC response theory [42]. Note that this expression is linear in the excited determinant components $l_{m\mu}$.

2. Ground-state left and excited-state right vectors

The left vector of the ground state has the components $l_{00} = 1$ and $l_{0\mu} = \bar{t}_\mu$ [see Eq. (19)], while the right vector of excited state n has the reference component $r_{0n} = -\sum_\nu \bar{t}_\nu r_{\nu n}$ [see Eq. (22)]. Inserting this into Eq. (A1), we obtain

$$ X_{0n} = \sum_\nu \left({}^{\mathrm{LR}}\eta_\nu + \sum_\mu \bar{t}_\mu \langle \mu | \tau_\nu \bar{X} | \mathrm{HF} \rangle \right) r_{\nu n}$$
$$ - \left(X_{00} + \sum_\mu \bar{t}_\mu \xi_\mu^X \right) \sum_\nu \bar{t}_\nu r_{\nu n}, \tag{A6}$$

where

$$ {}^{\mathrm{LR}}\eta_\nu = \langle \mathrm{HF} [\bar{X}, \tau_\nu] \mathrm{HF} \rangle + \sum_\mu \bar{t}_\mu {}^{\mathrm{LR}}A_{\mu\nu}^X. \tag{A7}$$

The term $\sum_\nu {}^{\mathrm{LR}}\eta_\nu r_{\nu n}$ appears in CC response theory [42], and the other terms are specific to EOM-CC. Equation (A6) is equivalent to Eq. (65) in Ref. [57] and can also be written as

$$ X_{0n} = \sum_\nu \left({}^{\mathrm{EOM}}\eta_\nu^X - X_{00}\bar{t}_\nu \right) r_{\nu n}, \tag{A8}$$

where

$$ {}^{\mathrm{EOM}}\eta_\nu^X = {}^{\mathrm{LR}}\eta_\nu^X + \sum_\mu \bar{t}_\mu \langle \mu | \tau_\nu \bar{X} | \mathrm{HF} \rangle - \left(\sum_\mu \bar{t}_\mu \xi_\mu^X \right) \bar{t}_\nu \tag{A9}$$

[see Eq. (18) of Ref. [58]]. Note that Eq. (A8) is linear in the excited determinant components $r_{\nu n}$.

3. Excited-state left and right vectors

The left vector of excited state m has the reference component $l_{m0} = 0$, while the right vector of excited state n has $r_{0n} = -\sum_\nu \bar{t}_\nu r_{\nu n}$ [see Eqs. (22) and (23)]. Inserting this into Eq. (A1), we obtain

$$ X_{mn} = \sum_{\mu\nu} l_{m\mu} \left({}^{\mathrm{LR}}A_{\mu\nu}^X + \langle \mu | \tau_\nu \bar{X} | \mathrm{HF} \rangle \right) r_{\nu n}$$
$$ - \sum_\mu l_{m\mu} \xi_\mu^X \sum_\nu \bar{t}_\nu r_{\nu n}$$
$$ = \sum_{\mu\nu} l_{m\mu} \left({}^{\mathrm{LR}}A_{\mu\nu}^X + \langle \mu | \tau_\nu \bar{X} | \mathrm{HF} \rangle - \xi_\mu^X \bar{t}_\nu \right) r_{\nu n}$$
$$ = \sum_{\mu\nu} l_{m\mu} \left({}^{\mathrm{EOM}}A_{\mu\nu}^X + \delta_{\mu\nu} X_{00} - \xi_\mu^X \bar{t}_\nu \right) r_{\nu n}, \tag{A10}$$

where

$$ {}^{\mathrm{EOM}}A_{\mu\nu}^X = \langle \mu | \bar{X} | \nu \rangle - \delta_{\mu\nu} X_{00}$$
$$ = {}^{\mathrm{LR}}A_{\mu\nu}^X + \langle \mu | \tau_\nu \bar{X} | \mathrm{HF} \rangle - \delta_{\mu\nu} X_{00} \tag{A11}$$

[see Eq. (20) in Ref. [58]].

The term $\sum_{\mu\nu} l_{m\mu} {}^{\mathrm{LR}}A_{\mu\nu}^X r_{\nu n}$ appears in CC response theory, and the other terms are specific to EOM-CC [44,58]. Note that all matrix elements in Eq. (A10) are linear in both $l_{m\mu}$ and $r_{\nu n}$.

APPENDIX B: INTEGRATION SCHEME

In order to limit the error of the time-dependent results, the integration of the TDCC and TD-EOM-CC equations is done with the embedded Dormand-Prince method of order 5(4) [53]. This method yields both fourth- and fifth-order accurate solutions at each time step, and is specified by the Butcher tableau [53,59]

$$
\begin{array}{c|ccccccc}
0 & & & & & & & \\
\frac{1}{5} & \frac{1}{5} & & & & & & \\
\frac{3}{10} & \frac{3}{40} & \frac{9}{40} & & & & & \\
\frac{4}{5} & \frac{44}{45} & -\frac{56}{15} & \frac{32}{9} & & & & \\
\frac{8}{9} & \frac{19372}{6561} & -\frac{25360}{2187} & \frac{64448}{6561} & -\frac{212}{729} & & & \\
1 & \frac{9017}{3168} & -\frac{355}{33} & \frac{46732}{5247} & \frac{49}{176} & -\frac{5103}{18656} & & \\
1 & \frac{35}{384} & 0 & \frac{500}{1113} & \frac{125}{192} & -\frac{2187}{6784} & \frac{11}{84} & \\
\hline
& \frac{35}{384} & 0 & \frac{500}{1113} & \frac{125}{192} & -\frac{2187}{6784} & \frac{11}{84} & 0 \\
& \frac{5179}{57600} & 0 & \frac{7571}{16695} & \frac{393}{640} & -\frac{92097}{339200} & \frac{187}{2100} & \frac{1}{40}
\end{array} \tag{B1}
$$

where the next-to-last and last rows give the coefficients of the fifth- and fourth-order solutions, respectively. Although the method has seven stages, its *first same as last* property assures that only six function evaluations are needed per time step.

The Euclidean distance between the solutions gives a fourth-order estimate of the local integration error:

$$ \epsilon_{O(4)} = ||\mathbf{y}_{O(5)} - \mathbf{y}_{O(4)}||_2. \tag{B2}$$

This local error estimate is kept below a given maximum value by adapting the time step during the integration. The fifth-order solution is accepted as the solution at the beginning of the next step whenever the error estimate satisfies this condition.

The following adaptive time-stepping scheme was designed, implemented, and used together with the Dormand-Prince 5(4) method for the relevant calculations in Sec. IV. At the start of the integration, the step size is set to a given maximum value. During the integration, the variable step size is halved, and the integration step redone, whenever the error estimate exceeds the given maximum error. After a successful integration step, the step size is doubled whenever the error estimate is below a given minimum value, provided that the doubled step size is smaller than the maximum step size and also a submultiple of the elapsed time. This is in order to increase the efficiency of the integration while ensuring that the solution is evaluated at times corresponding to integer increments of the maximum time-step size. Evaluation of time-dependent observables is done using the solutions at these integer increments.

[1] M. T. Hassan, T. T. Luu, A. Moulet, O. Raskazovskaya, P. Zhokhov, M. Garg, N. Karpowicz, A. M. Zheltikov, V. Pervak, F. Krausz, and E. Goulielmakis, Nature (London) **530**, 66 (2016).

[2] M. Galli, V. Wanie, D. P. Lopes, E. P. Månsson, A. Trabattoni, L. Colaizzi, K. Saraswathula, A. Cartella, F. Frassetto, L. Poletto, F. Légaré, S. Stagira, M. Nisoli, R. M. Vázquez, R. Osellame, and F. Calegari, Opt. Lett. **44**, 1308 (2019).

[3] D. Fabris, T. Witting, W. A. Okell, D. J. Walke, P. Matia-Hernando, J. Henkel, T. R. Barillot, M. Lein, J. P. Marangos, and J. W. G. Tisch, Nat. Photon. **9**, 383 (2015).

[4] J. Duris, S. Li, T. Driver, E. G. Champenois, J. P. MacArthur, A. A. Lutman, Z. Zhang, P. Rosenberger, J. W. Aldrich, R. Coffee, G. Coslovich, F.-J. Decker, J. M. Glownia, G. Hartmann, W. Helml, A. Kamalov, J. Knurr, J. Krzywinski, M.-F. Lin, J. P. Marangos *et al.*, Nat. Photon. **14**, 30 (2020).

[5] A. Malyzhenkov, Y. P. Arbelo, P. Craievich, P. Dijkstal, E. Ferrari, S. Reiche, T. Schietinger, P. Juranić, and E. Prat, Phys. Rev. Research **2**, 042018(R) (2020).

[6] X. Li, N. Govind, C. Isborn, A. E. DePrince, and K. Lopata, Chem. Rev. **120**, 9951 (2020).

[7] B. F. E. Curchod and T. J. Martínez, Chem. Rev. **118**, 3305 (2018).

[8] E. F. Kjønstad and H. Koch, J. Chem. Theory Comput. **17**, 127 (2021).

[9] S. E. Koonin, K. T. R. Davies, V. Maruhn-Rezwani, H. Feldmeier, S. J. Krieger, and J. W. Negele, Phys. Rev. C **15**, 1359 (1977).

[10] P. Bonche, B. Grammaticos, and S. Koonin, Phys. Rev. C **17**, 1700 (1978).

[11] H. Flocard, S. E. Koonin, and M. S. Weiss, Phys. Rev. C **17**, 1682 (1978).

[12] M. R. Provorse and C. M. Isborn, Int. J. Quantum Chem. **116**, 739 (2016).

[13] C. Lian, M. Guan, S. Hu, J. Zhang, and S. Meng, Adv. Theory Simul. **1**, 1800055 (2018).

[14] H. Miyagi and L. B. Madsen, Phys. Rev. A **87**, 062511 (2013).

[15] T. Sato and K. L. Ishikawa, Phys. Rev. A **88**, 023402 (2013).

[16] H. Miyagi and L. B. Madsen, Phys. Rev. A **89**, 063416 (2014).

[17] P. Krause, T. Klamroth, and P. Saalfrank, J. Chem. Phys. **123**, 074105 (2005).

[18] J. A. Sonk, M. Caricato, and H. B. Schlegel, J. Phys. Chem. A **115**, 4678 (2011).

[19] E. Luppi and M. Head-Gordon, Mol. Phys. **110**, 909 (2012).

[20] A. E. DePrince, M. Pelton, J. R. Guest, and S. K. Gray, Phys. Rev. Lett. **107**, 196806 (2011).

[21] L. Cederbaum and J. Zobeley, Chem. Phys. Lett. **307**, 205 (1999).

[22] M. Ruberti, Faraday Discuss. **228**, 286 (2021).

[23] D. R. Nascimento and A. E. DePrince, J. Phys. Chem. Lett. **8**, 2951 (2017).

[24] L. N. Koulias, D. B. Williams-Young, D. R. Nascimento, A. E. DePrince, and X. Li, J. Chem. Theory Comput. **15**, 6617 (2019).

[25] S. Kvaal, J. Chem. Phys. **136**, 194109 (2012).

[26] C. Huber and T. Klamroth, J. Chem. Phys. **134**, 054113 (2011).

[27] D. A. Pigg, G. Hagen, H. Nam, and T. Papenbrock, Phys. Rev. C **86**, 014308 (2012).

[28] A. S. Skeidsvoll, A. Balbi, and H. Koch, Phys. Rev. A **102**, 023115 (2020).

[29] Y. C. Park, A. Perera, and R. J. Bartlett, J. Chem. Phys. **151**, 164117 (2019).

[30] T. B. Pedersen and S. Kvaal, J. Chem. Phys. **150**, 144106 (2019).

[31] H. E. Kristiansen, Ø. S. Schøyen, S. Kvaal, and T. B. Pedersen, J. Chem. Phys. **152**, 071102 (2020).

[32] T. B. Pedersen, H. E. Kristiansen, T. Bodenstein, S. Kvaal, and Ø. S. Schøyen, J. Chem. Theory Comput. **17**, 388 (2021).

[33] Y. C. Park, A. Perera, and R. J. Bartlett, J. Chem. Phys. **155**, 094103 (2021).

[34] M. Wu, S. Chen, S. Camp, K. J. Schafer, and M. B. Gaarde, J. Phys. B **49**, 062003 (2016).

[35] T. Chen, J. Demmel, M. Gu, Y. Saad, R. Lehoucq, D. Sorensen, K. Maschhoff, Z. Bai, D. Day, R. Freund, G. Sleijpen, H. van der Vorst, and R. Li, in *Templates for the Solution of Algebraic Eigenvalue Problems* (SIAM, Philadelphia, 2000), Chap. 7, pp. 149–231.

[36] I. H. Godtliebsen and O. Christiansen, Phys. Chem. Chem. Phys. **15**, 10035 (2013).

[37] R. W. Freund, Acta Numer. **12**, 267 (2003).

[38] N. Shimizu, T. Mizusaki, Y. Utsuno, and Y. Tsunoda, Comput. Phys. Commun. **244**, 372 (2019).

[39] J. K. Cullum and R. A. Willoughby, *Lanczos Algorithms for Large Symmetric Eigenvalue Computations, Vol. I: Theory* (SIAM, Philadelphia, 2002).

[40] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 4th ed. (Johns Hopkins University, Baltimore, MD, 2013).

[41] Y. Saad, *Iterative Methods for Sparse Linear Systems* (SIAM, Philadelphia, 2003).

[42] S. Coriani, T. Fransson, O. Christiansen, and P. Norman, J. Chem. Theory Comput. **8**, 1616 (2012).

[43] S. Coriani, O. Christiansen, T. Fransson, and P. Norman, Phys. Rev. A **85**, 022507 (2012).

[44] B. N. C. Tenorio, M. A. C. Nascimento, A. B. Rocha, and S. Coriani, J. Chem. Phys. **151**, 184106 (2019).

[45] T. Moitra, S. Coriani, and B. N. C. Tenorio, Mol. Phys. **119**, e1980235 (2021).

[46] NIST Computational Chemistry Comparison and Benchmark Database, NIST Standard Reference Database No. 101 (Release 20 August 2019).

[47] R. A. Kendall, T. H. Dunning, and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).

[48] D. E. Woon and T. H. Dunning, J. Chem. Phys. **103**, 4572 (1995).

[49] L. S. Cederbaum, W. Domcke, and J. Schirmer, Phys. Rev. A **22**, 206 (1980).

[50] S. Coriani and H. Koch, J. Chem. Phys. **143**, 181103 (2015).

[51] B. N. C. Tenorio, T. Moitra, M. A. C. Nascimento, A. B. Rocha, and S. Coriani, J. Chem. Phys. **150**, 224104 (2019).

[52] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevA.105.023103 for the energies and maximum transition strengths for all EOM-CC states used in the TD-EOM-CC calculations in Sec. IV.

[53] J. Dormand and P. Prince, J. Comput. Appl. Math. **6**, 19 (1980).

[54] S. D. Folkestad, E. F. Kjønstad, R. H. Myhre, J. H. Andersen, A. Balbi, S. Coriani, T. Giovannini, L. Goletto, T. S. Haugland, A. Hutcheson, I.-M. Høyvik, T. Moitra, A. C. Paul, M. Scavino, A. S. Skeidsvoll, Å. H. Tveten, and H. Koch, J. Chem. Phys. **152**, 184103 (2020).

[55] A. D. Dutoi, K. Gokhberg, and L. S. Cederbaum, Phys. Rev. A **88**, 013419 (2013).

[56] A. D. Dutoi and L. S. Cederbaum, Phys. Rev. A **90**, 023414 (2014).

[57] H. Koch, R. Kobayashi, A. Sanchez de Merás, and P. Jørgensen, J. Chem. Phys. **100**, 4393 (1994).

[58] R. Faber and S. Coriani, J. Chem. Theory Comput. **15**, 520 (2019).

[59] A. Iserles, in *A First Course in the Numerical Analysis of Differential Equations*, 2nd ed., Cambridge Texts in Applied Mathematics (Cambridge University Press, New York, 2008), pp. 105–122.

Paper IV

Comparing real-time coupled cluster methods through simulation of collective Rabi oscillations

Andreas S. Skeidsvoll

Department of Chemistry, Norwegian University of Science and Technology, 7491 Trondheim, Norway

Henrik Koch*

Scuola Normale Superiore, Piazza dei Cavalieri, 7, I-56126, Pisa, Italy   and
Department of Chemistry, Norwegian University of Science and Technology, 7491 Trondheim, Norway
(Dated: May 23, 2023)

The time-dependent equation-of-motion coupled cluster singles and doubles (TD-EOM-CCSD) and time-dependent coupled cluster singles and doubles (TDCCSD) methods are compared by simulating Rabi oscillations for distant atoms in a classical electromagnetic field. While the TD-EOM-CCSD simulations are numerically stable, the oscillating time-dependent energy expectation value displays an incorrect scaling with the number of atoms resonant with the field. On the other hand, the TDCCSD simulations exhibit the correct scaling in the initial stages of the Rabi cycle, but break down when the multi-atom system approaches complete population inversion. The numerical stability of the TDCCSD simulations is shown to be very sensitive to any interactions between the atoms, and to left amplitudes that do not participate in the expression for time-dependent expectation values for non-interacting atoms. Moreover, we introduce a general theoretical framework for describing the two methods, where the cluster amplitude time derivatives serve as auxiliary conditions related to a shift of the time-dependent Hamiltonian matrix. In this framework, time-dependent coupled cluster methods exhibit a shifted Hamiltonian matrix with a block upper triangular structure in terms of the number of excited non-interacting subsystems, explaining the correct scaling properties of TDCCSD.

I. INTRODUCTION

With recent developments in the shaping and amplification of laser pulses, the production of short and strong pulses can now be realized in several frequency domains [1–3]. The progress sparks further interest in the dynamical and non-linear response of molecules to strong fields, which can involve a high number of quantum states [4]. This since many of the states that are inaccessible by a single-photon transition, either energetically or by symmetry selection rules, can be accessed by a multiphoton transition [5, 6].

The ultrafast non-linear response of molecules to strong fields can give an extended degree of dynamic control of chemical reactions [7–9]. It can also reveal information about the system that is inaccessible in weaker fields, which can be used for improving the imaging of different reaction stages [5, 8]. That said, the involved coupling between the numerous affected states can lead to an intricate relationship between the shape and strength of laser pulses and the molecular response, which calls for the interpretation by appropriate quantum chemistry methods [10].

Both the accuracy and computational complexity of quantum chemistry methods often increase with the order of approximation of the particle correlations in the system [10, 11] and the size of the finite basis set [12], but the accuracy also depends heavily on the mathematical

structure of the method. Considerable effort is spent on constructing the most well-behaved methods for a given order of computational complexity, with respect to both numerical stability and correspondence to experimental results for various systems. Real-time variants of quantum chemistry methods are convenient for modeling multiphoton transitions in systems [13], as expressions for the high-order frequency response can be difficult to both derive and to solve numerically.

The well-established single-reference coupled cluster (CC) hierarchy of methods often gives accurate and rapidly converging molecular properties [14] for states with weak multi-reference character [15]. An important reason for this accuracy is the physically reasonable scaling properties of the methods, even when the cluster operator is truncated. For instance, the energy of the ground state is size-extensive, meaning that it scales linearly with the number of non-interacting identical subsystems [16]. In the equation-of-motion (EOM-CC) framework, the excitation energies are size-intensive, meaning that they do not scale with the number of non-interacting subsystems [17]. In the linear response framework, which is based on time-dependent coupled cluster theory, ground state—excited state transition moments are also size-intensive [18]. Truncated configuration interaction methods, on the other hand, do not possess these properties, and errors generally increase with the size of the simulated system [19].

Traditionally, the coupled cluster methods have almost exclusively been treated in the frequency domain, but the last decade has witnessed an increased exploration into their real-time behavior [13]. As demonstrated by

Pedersen and Kvaal [20] and further investigated by Kristiansen et al. [21], the exponential parametrization makes the standard time-dependent coupled cluster (TDCC) method inherently unstable whenever the reference determinant weight is depleted by a strong field. These instabilities can require the use of exceedingly small time steps in numerical solutions, and can at times also lead to breakdowns that cannot be solved by decreasing the time step size.

The orbital adaptive time-dependent coupled cluster (OATDCC) method, which requires the solution of an additional set of linear equations at each time step, was shown to be more stable than TDCC. Nonetheless, the method still fails at higher field strengths, where the reference determinant weight can become greater than one [21].

Variants of the time-dependent equation-of-motion coupled cluster (TD-EOM-CC) method have also been used for modeling laser-molecule interactions, but only a handful applications have included the full non-linear real-time propagation of the laser-driven electron dynamics [13, 22–24]. In these cases, the TD-EOM-CC equations were expressed in the basis obtained by diagonalizing the field-free equation-of-motion coupled cluster Hamiltonian. We instead express the equations in the elementary basis, leading to equations that are simple to implement and have computational and memory requirements that scale more favorably with respect to system size than the full diagonal basis equations. This makes the formulation particularly useful for assessing the short-time and non-linear behavior of the TD-EOM-CC method.

The paper is organized as follows. In Section II, the TD-EOM-CC and TDCC methods are described in a general theoretical framework, and it is shown how the time derivative of the cluster amplitudes affects the analytical scaling properties of the truncated variants of the methods. Section III outlines the computational methods used to simulate distant atoms undergoing semiclassical Rabi oscillations in a resonant electromagnetic field. In Section IV, results of the simulations are presented and discussed, including a demonstration of how the time-dependent energy expectation value scales with respect to system size in TD-EOM-CCSD and TDCCSD. The key findings are summarized in Section V.

II. THEORY

A. System

The time-dependent system of the molecule and the electromagnetic field is described by the Hamiltonian

$$H(t) = H^{(0)} + V(t). \qquad (1)$$

The field-free molecular system is described by the Hamiltonian $H^{(0)}$, and the interaction between the molecular system and the electromagnetic field is described by $V(t)$. We describe the interaction semiclassically, in the dipole approximation and length gauge. This gives $V(t) = -\boldsymbol{\mu} \cdot \boldsymbol{\mathcal{E}}(t)$, where $\boldsymbol{\mu}$ is the electric dipole moment vector and $\boldsymbol{\mathcal{E}}(t)$ the classical time-dependent electric field vector. The system is also treated within the Born-Oppenheimer approximation, with fixed nuclei.

B. Time dependence in coupled cluster methods

The TD-EOM-CC and TDCC methods can be written in a general theoretical framework by expressing the time-dependent ket and bra vectors as

$$|\Psi(t)\rangle = e^{T(t)} R(t) |\text{HF}\rangle, \qquad (2)$$

$$\langle \widetilde{\Psi}(t)| = \langle \text{HF}| L(t) e^{-T(t)}, \qquad (3)$$

where the cluster operator

$$T(t) = \sum_{\kappa \geq 0} \tau_\kappa t_\kappa(t) \qquad (4)$$

and the right and left operators

$$R(t) = \sum_{\kappa \geq 0} \tau_\kappa r_\kappa(t), \quad L(t) = \sum_{\kappa \geq 0} l_\kappa(t) \widetilde{\tau}_\kappa^\dagger. \qquad (5)$$

The operators τ_0 and $\widetilde{\tau}_0^\dagger$ are the unit operator,

$$\tau_0 = \widetilde{\tau}_0^\dagger = \mathbb{1}, \qquad (6)$$

and the operators τ_μ and $\widetilde{\tau}_\mu^\dagger$, where $\mu > 0$, excite and deexcite electrons between occupied and virtual Hartree-Fock molecular orbitals, respectively,

$$\tau_\mu |\text{HF}\rangle = |\mu\rangle, \quad \langle \text{HF}| \widetilde{\tau}_\mu^\dagger = \langle \widetilde{\mu}|, \qquad (7)$$

$$\widetilde{\tau}_\mu^\dagger |\text{HF}\rangle = 0, \quad \langle \text{HF}| \tau_\mu = 0. \qquad (8)$$

These operators are chosen so that the resulting determinants are biorthonormal,

$$\langle \widetilde{\kappa}|\lambda\rangle = \delta_{\kappa\lambda}, \quad \kappa \geq 0, \quad \lambda \geq 0, \qquad (9)$$

where $\delta_{\kappa\lambda}$ is the Kronecker delta. The sums over excitation levels in Eq. (4) and Eq. (5) can be truncated in order to reduce computational scaling, yielding the coupled cluster singles method (CCS) when only single excitations are included, coupled cluster singles and doubles (CCSD) when both single and double excitations are included, and so on.

The field-free coupled cluster ground state can be defined by setting the right and left amplitudes $r_\mu^{(0)} = 0$ and $r_0^{(0)} = l_0^{(0)} = 1$, and letting the ground state cluster amplitudes $t_\mu^{(0)}$ and left amplitudes $l_\mu^{(0)}$ be determined as solutions of the field-free ground state equations

$$\langle \widetilde{\mu}| e^{-T^{(0)}} H^{(0)} e^{T^{(0)}} |\text{HF}\rangle = 0, \qquad (10)$$

$$\left(\langle \text{HF}| + \sum_{\mu > 0} l_\mu^{(0)} \langle \widetilde{\mu}| \right) [e^{-T^{(0)}} H^{(0)} e^{T^{(0)}}, \tau_\nu] |\text{HF}\rangle = 0. \qquad (11)$$

For simplicity, we also set the undetermined phase-related cluster amplitude $t_0^{(0)} = 0$.

The equations for the time dependence of the parameters of Eq. (5) can be derived from the right and left time-dependent Schrödinger equations (TDSEs)

$$i\frac{\mathrm{d}}{\mathrm{d}t}\left|\Psi(t)\right\rangle = H(t)\left|\Psi(t)\right\rangle, \tag{12}$$

$$-i\frac{\mathrm{d}}{\mathrm{d}t}\left\langle\widetilde{\Psi}(t)\right| = \left\langle\widetilde{\Psi}(t)\right| H(t). \tag{13}$$

Inserting Eq. (2) into Eq. (12) before projecting onto $\langle\widetilde{\kappa}|\,e^{-T(t)}$, and likewise inserting Eq. (3) into Eq. (13) before projecting onto $e^{T(t)}\left|\lambda\right\rangle$, the following matrix-vector TDSEs are obtained

$$i\frac{\mathrm{d}r_\kappa(t)}{\mathrm{d}t} = \sum_{\lambda\geq 0}\widetilde{H}_{\kappa\lambda}(t)r_\lambda(t), \tag{14}$$

$$-i\frac{\mathrm{d}l_\lambda(t)}{\mathrm{d}t} = \sum_{\kappa\geq 0}l_\kappa(t)\widetilde{H}_{\kappa\lambda}(t), \tag{15}$$

where the shifted Hamiltonian

$$\widetilde{H}(t) = H(t) - i\frac{\mathrm{d}T(t)}{\mathrm{d}t}. \tag{16}$$

The elements of the coupled cluster matrix $\boldsymbol{O}(t)$ of operator $O(t)$ are given by

$$O_{\kappa\lambda}(t) = \langle\widetilde{\kappa}|\,\overline{O}(t)\,|\lambda\rangle, \tag{17}$$

where an overbar is used to denote the similarity transformation by the exponentiated time-dependent cluster operator,

$$\overline{O}(t) = e^{-T(t)}O(t)e^{T(t)}. \tag{18}$$

By invoking the resolution of identity $\mathbb{1} = |\mathrm{HF}\rangle\langle\mathrm{HF}| + \sum_{\mu>0}|\mu\rangle\langle\widetilde{\mu}|$, the matrix elements of the shifted Hamiltonian in Eq. (16) can be expressed as

$$\begin{aligned}
\widetilde{H}_{\kappa\lambda}(t) &= \langle\widetilde{\kappa}|\,\overline{H}(t)\,|\lambda\rangle - i\langle\widetilde{\kappa}|\left(\sum_{\eta\geq 0}\tau_\eta\frac{\mathrm{d}t_\eta(t)}{\mathrm{d}t}\right)|\lambda\rangle \\
&= \langle\widetilde{\kappa}|\,[\overline{H}(t),\tau_\lambda]\,|\mathrm{HF}\rangle \\
&\quad + \langle\widetilde{\kappa}|\,\tau_\lambda\left(|\mathrm{HF}\rangle\langle\mathrm{HF}| + \sum_{\mu>0}|\mu\rangle\langle\widetilde{\mu}|\right)\overline{H}(t)\,|\mathrm{HF}\rangle \\
&\quad - i\langle\widetilde{\kappa}|\lambda\rangle\frac{\mathrm{d}t_0(t)}{\mathrm{d}t} - i\sum_{\mu>0}\langle\widetilde{\kappa}|\,\tau_\mu\,|\lambda\rangle\frac{\mathrm{d}t_\mu(t)}{\mathrm{d}t} \\
&= U_{\kappa\lambda}(t) + \widetilde{D}_{\kappa\lambda}(t) + \widetilde{L}_{\kappa\lambda}(t),
\end{aligned} \tag{19}$$

where

$$U_{\kappa\lambda}(t) = \langle\widetilde{\kappa}|\,[\overline{H}(t),\tau_\lambda]\,|\mathrm{HF}\rangle, \tag{20}$$

$$\widetilde{D}_{\kappa\lambda}(t) = \delta_{\kappa\lambda}\left(\langle\mathrm{HF}|\,\overline{H}(t)\,|\mathrm{HF}\rangle - i\frac{\mathrm{d}t_0(t)}{\mathrm{d}t}\right), \tag{21}$$

$$\widetilde{L}_{\kappa\lambda}(t) = \sum_{\mu>0}\langle\widetilde{\kappa}|\,\tau_\mu\,|\lambda\rangle\left(\langle\mu|\,\overline{H}(t)\,|\mathrm{HF}\rangle - i\frac{\mathrm{d}t_\mu(t)}{\mathrm{d}t}\right). \tag{22}$$

To explain the choices of the names of the three terms in Eq. (19) in anticipation of the discussion in II C, we remark that the commutator $[\overline{H}(t),\tau_\lambda]$ in Eq. (20) implies that the $\boldsymbol{U}(t)$ term of the shifted Hamiltonian matrix for two non-interacting subsystems is block upper triangular in terms of the number of excited subsystems. Furthermore, the excitation operator τ_μ in Eq. (22), implies that the $\boldsymbol{L}(t)$ term is block lower triangular, and the Kronecker delta in Eq. (21) implies that the $\boldsymbol{D}(t)$ term is (block) diagonal.

Once all time-dependent amplitudes have been found at a given point in time t, the time-dependent expectation values of the time-dependent operator $O(t)$ can be obtained through

$$\begin{aligned}
\langle O(t)\rangle &= \langle\widetilde{\Psi}(t)|\,O(t)\,|\Psi(t)\rangle \\
&= \sum_{\kappa,\lambda\geq 0}l_\kappa(t)O_{\kappa\lambda}(t)r_\lambda(t) \\
&= \boldsymbol{l}^T(t)\boldsymbol{O}(t)\boldsymbol{r}(t).
\end{aligned} \tag{23}$$

The indeterminate matrix-vector equations in Eq. (14) and Eq. (15) give a unified representation of the TD-EOM-CC and TDCC methods, since the two methods can be recovered by imposing constraints on the cluster amplitude time derivatives and assuming that the propagation starts from the ground state. The TD-EOM-CC method is usually derived without consideration for the time dependence of the cluster amplitudes [22–24], however the time-dependence can equivalently be removed after the derivation of Eq. (14) and Eq. (15) by setting

$$i\frac{\mathrm{d}t_\kappa(t)}{\mathrm{d}t} = 0. \tag{24}$$

Since we also assume that the propagation starts from the ground state, Eq. (24) implies that $T(t) = T^{(0)}$, and Eq. (19) can thus be expressed as

$$\widetilde{H}_{\kappa\lambda}(t) = U_{\kappa\lambda}(t) + \widetilde{D}_{\kappa\lambda}(t) + \widetilde{L}_{\kappa\lambda}(t), \tag{25}$$

where

$$U_{\kappa\lambda}(t) = \langle\widetilde{\kappa}|\,[e^{-T^{(0)}}H(t)e^{T^{(0)}},\tau_\lambda]\,|\mathrm{HF}\rangle, \tag{26}$$

$$\widetilde{D}_{\kappa\lambda}(t) = \delta_{\kappa\lambda}\,\langle\mathrm{HF}|\,e^{-T^{(0)}}H(t)e^{T^{(0)}}\,|\mathrm{HF}\rangle, \tag{27}$$

$$\widetilde{L}_{\kappa\lambda}(t) = \sum_{\mu>0}\langle\widetilde{\kappa}|\,\tau_\mu\,|\lambda\rangle\,\langle\mu|\,e^{-T^{(0)}}V(t)e^{T^{(0)}}\,|\mathrm{HF}\rangle, \tag{28}$$

and the $H^{(0)}$ term of $\widetilde{L}_{\kappa\lambda}(t)$ is zero because of Eq. (10).

The TDCC time derivatives of the phase $\epsilon(t)$ and cluster amplitudes $t_\mu(t)$ [25] can be gathered in the expression

$$i\frac{\mathrm{d}t_\kappa(t)}{\mathrm{d}t} = \langle\widetilde{\kappa}|\,\overline{H}(t)\,|\mathrm{HF}\rangle \tag{29}$$

by making the identification $\epsilon(t) = -it_0(t)$. From this expression, Eq. (21) and Eq. (22) can be seen to reduce to

$$\widetilde{D}_{\kappa\lambda}(t) = 0, \tag{30}$$

$$\widetilde{L}_{\kappa\lambda}(t) = 0, \tag{31}$$

implying that

$$\begin{aligned}
\widetilde{H}_{\kappa\lambda}(t) &= U_{\kappa\lambda}(t) \\
&= \langle\widetilde{\kappa}|\left[\overline{H}(t),\tau_\lambda\right]|\mathrm{HF}\rangle
\end{aligned} \tag{32}$$

in TDCC.

The ket and bra vectors used in standard derivations of the TDCC method [25] do not include the right operator $R(t)$ in Eq. (2), nor the time-dependence of the reference component of $L(t)$ in Eq. (3). To demonstrate that the TDCC method is still retrieved when the propagation starts from the ground state, we start by noting that Eq. (14) and Eq. (32) imply that the time derivatives of all $r_\kappa(t)$ are zero whenever all $r_\mu(t)$ are zero for $\mu > 0$. Furthermore, Eq. (15) and Eq. (32) imply that the time derivative of $l_0(t)$ is zero. Therefore, when the propagation starts from the coupled cluster ground state, $r_\mu(t) = 0$ and $l_0(t) = r_0(t) = 1$ at all times t. For $\kappa > 0$, Eq. (15) is hence equivalent to the expression for the left amplitude time derivatives in TDCC [25], and the TDCC method is retrieved despite the additional flexibility in the parametrization of Eq. (2) and Eq. (3).

The choice of cluster amplitude time derivatives has major implications for the properties of the resulting method. In TD-EOM-CC, Eq. (24) implies that the cluster amplitudes do not change with time, and real arithmetic can be used to reduce the cost of TD-EOM-CC propagation provided the cluster amplitudes of the ground state are real. In TDCC, on the hand, Eq. (29) implies that the time-dependent cluster amplitudes are complex, requiring the use of complex arithmetic. The TDCC time derivatives also imply that Eq. (14) and Eq. (15) are non-linear in the time-dependent parameters, and the expression and interpretation of the time-dependent state in terms of EOM-CC states is thus nontrivial [26]. For TD-EOM-CC, on the other hand, Eq. (14) and Eq. (15) are linear in the time-dependent parameters, and the time independence of the cluster operator allows for the method to be expressed in the basis of the field-free EOM-CC states. This can be shown by operating with the resolution of identity

$$\mathbb{1} = \sum_{i\geq 0}|\psi_i^{(0)}\rangle\,\langle\widetilde{\psi}_i^{(0)}| \tag{33}$$

onto both Eq. (2) and Eq. (3), where the ket and bra of field-free EOM-CC state i can be expressed as

$$|\psi_i^{(0)}\rangle = R_i^{(0)}e^{T^{(0)}}|\mathrm{HF}\rangle, \quad \langle\psi_i^{(0)}| = \langle\mathrm{HF}|\,e^{T^{(0)}}L_i^{(0)}, \tag{34}$$

and the amplitudes of the operators $R_i^{(0)} = \sum_{\kappa\geq 0}\tau_\kappa R_{\kappa i}^{(0)}$ and $L_i^{(0)} = \sum_{\kappa\geq 0}L_{i\kappa}^{(0)}\tau_\kappa^\dagger$ are the right and left eigenvectors of the field-free coupled cluster Hamiltonian matrix $H_{\kappa\lambda}^{(0)} = \langle\kappa|\,e^{-T^{(0)}}H^{(0)}e^{T^{(0)}}|\lambda\rangle$. By letting $\langle\widetilde{\psi}_i^{(0)}|\Psi(t)\rangle = s_i(t)$ and $\langle\widetilde{\Psi}(t)|\psi_i^{(0)}\rangle = k_i(t)$, the elementary basis formulation of TD-EOM-CC presented here can be seen to be equivalent to the diagonal basis formulation of Ref. [24]. As demonstrated in that reference, field-free EOM-CC

states that do not participate in the dynamics can be removed from the diagonal basis in order to reduce computational costs. However, the full elementary basis formulation of TD-EOM-CC, as given by Eq. (14), Eq. (15) and Eq. (25), straightforwardly ensures the inclusion of all transitions without requiring the calculation of all field-free EOM-CC states and transition moments.

C. Scaling properties of real-time coupled cluster methods

In order to theoretically investigate the scaling properties of methods based on the parametrization in Eq. (2) and Eq. (3), we assume that the system is composed of multiple non-interacting subsystems. We let $\tau_{\lambda_\mathrm{I}}$ denote an elementary excitation operator and $\widetilde{\tau}_{\kappa_\mathrm{I}}^\dagger$ an elementary deexcitation operator of subsystem I. The elementary excitation and deexcitation operators of the composite system can be constructed as tensor products of all the operators of the different subsystems. Untruncated TD-EOM-CC and TDCC methods can represent all tensor products, since the excitation and deexcitation levels of the methods are not limited. In truncated methods, however, all elementary excitation and deexcitation operators that exceed the truncation level specific to the method are excluded, which can lead to errors related to the scaling from one to multiple subsystems.

For two subsystems $\mathrm{I} \in \{\mathrm{A}, \mathrm{B}\}$, the elementary excitation and deexcitation operators of the composite system can be constructed as the tensor products $\tau_{\lambda_\mathrm{A}} \otimes \tau_{\lambda_\mathrm{B}}$ and $\widetilde{\tau}_{\kappa_\mathrm{A}}^\dagger \otimes \widetilde{\tau}_{\kappa_\mathrm{B}}^\dagger$. We split the sets of these operators into four partitions, which we label by 0, A, B and AB. The 0 partition includes the operators that do not change the excitation level of the subsystems, $\tau_{0_\mathrm{A}} \otimes \tau_{0_\mathrm{B}}$ and $\widetilde{\tau}_{0_\mathrm{A}}^\dagger \otimes \widetilde{\tau}_{0_\mathrm{B}}^\dagger$. The A partition includes the operators that change the excitation level of subsystem A only, $\tau_{\mu_\mathrm{A}} \otimes \tau_{0_\mathrm{B}}$ and $\widetilde{\tau}_{\mu_\mathrm{A}}^\dagger \otimes \widetilde{\tau}_{0_\mathrm{B}}^\dagger$, and the B partition the operators that change the excitation level of subsystem B only, $\tau_{0_\mathrm{A}} \otimes \tau_{\mu_\mathrm{B}}$ and $\tau_{0_\mathrm{A}}^\dagger \otimes \tau_{\mu_\mathrm{B}}^\dagger$, where $\mu > 0$. The AB partition includes the operators that change the excitation level of both subsystems, $\tau_{\nu_\mathrm{A}} \otimes \tau_{\nu_\mathrm{B}}$ and $\widetilde{\tau}_{\mu_\mathrm{A}}^\dagger \otimes \widetilde{\tau}_{\mu_\mathrm{B}}^\dagger$, where $\mu > 0$ and $\nu > 0$. Truncation can affect the AB partition, since the tensor products of the truncated subsystem operators can include excitations and deexcitations that in combination go beyond the truncation level of the method. In the following, we assess the general impact of this truncation on the TD-EOM-CC and TDCC methods, without limiting the discussion to any particular truncation level.

We start by assuming that the cluster amplitudes corresponding to the operators $\tau_{\nu_\mathrm{A}} \otimes \tau_{\nu_\mathrm{B}}$ are zero at a given time t. The cluster operator $T(t)$ can then be written as the tensor sum,

$$T(t) = T_\mathrm{A}(t) \otimes I_\mathrm{B} + I_\mathrm{A} \otimes T_\mathrm{B}(t), \tag{35}$$

where $T_\mathrm{I}(t)$ is the cluster operator for subsystems I. Since operators on non-interacting subsystems commute, we

have that

$$e^{\pm\left(T_A(t)\otimes I_B + I_A\otimes T_B(t)\right)} = e^{\pm T_A(t)}\otimes e^{\pm T_B(t)}. \quad (36)$$

We furthermore let $O(t)$ be any operator that does not involve any interaction between the two subsystems and thus can be written as the tensor sum

$$O(t) = O_A(t)\otimes I_B + I_A\otimes O_B(t), \quad (37)$$

where $O_A(t)$ and $O_B(t)$ are subsystem operators. Equation (36) then implies that the similarity transformed operator in Eq. (18) can be written as the tensor sum

$$\begin{aligned}\overline{O}(t) &= e^{-T_A(t)}O_A(t)e^{T_A(t)}\otimes I_B \\ &+ I_A\otimes e^{-T_B(t)}O_B(t)e^{T_B(t)} \\ &= \overline{O}_A(t)\otimes I_B + I_A\otimes \overline{O}_B(t),\end{aligned} \quad (38)$$

which does not contain terms where both two subsystems are excited simultaneously.

We furthermore assume that the time-dependent Hamiltonian $H(t)$ does not involve any interaction between the two subsystems, implying that it can be written on the form of Eq. (37). In TDCC, the time derivative of the cluster amplitudes in Eq. (29) can for the AB partition then be written as

$$\begin{aligned}i\frac{dt_{\mu_A\mu_B}(t)}{dt} &= \big(\langle\widetilde{\mu}_A|\otimes\langle\widetilde{\mu}_B|\big) \\ &\times \big(\overline{H}_A(t)\otimes I_B + I_A\otimes\overline{H}_B(t)\big) \\ &\times \big(|HF_A\rangle\otimes|HF_B\rangle\big) \\ &= 0.\end{aligned} \quad (39)$$

If Eq. (35) is true at the initial time, it will remain true for later times in TDCC provided the Hamiltonian still separates as $H_A(t)\otimes I_B + I_A\otimes H_B(t)$. In TD-EOM-CC, the time derivatives of these cluster amplitudes are also zero as seen by Eq. (24).

We use the subscript $\parallel$ to denote that all amplitudes and matrix elements corresponding to the excitation and deexcitation operators that are not represented in truncated methods have been set to zero. This is used to formulate the truncated methods in the untruncated product basis, for a more straightforward comparison to untruncated methods. Accordingly, the truncated right and transposed left amplitude vectors $r_\parallel$ and $l_\parallel^T$ are obtained by eliminating components from the untruncated r and l^T, and can in the partitioned product bases be written as

$$r_\parallel = \begin{pmatrix} r_0 \\ r_A \\ r_B \\ (r_{AB})_\parallel \end{pmatrix}, \quad l_\parallel^T = \big(l_0\ \ l_A\ \ l_B\ \ (l_{AB})_\parallel\big), \quad (40)$$

where only the AB partitions are affected by the truncation. Moreover, the truncated operator matrix $O_\parallel$ is obtained from O, which is the projection of Eq. (38) onto

the kets $\tau_{\lambda_A}|HF_A\rangle\otimes\tau_{\lambda_B}|HF_B\rangle$ and bras $\langle HF_A|\widetilde{\tau}_{\kappa_A}^\dagger\otimes\langle HF_B|\widetilde{\tau}_{\kappa_B}^\dagger$. In the partitioned product bases, this matrix can be written as

$$O_\parallel = \begin{pmatrix} O_{00} & O_{0A} & O_{0B} & 0 \\ O_{A0} & O_{AA} & 0 & (O_{AAB})_\parallel \\ O_{B0} & 0 & O_{BB} & (O_{BAB})_\parallel \\ 0 & (O_{ABA})_\parallel & (O_{ABB})_\parallel & (O_{ABAB})_\parallel \end{pmatrix}, \quad (41)$$

where only the AB partition row and column are affected by the truncation.

We proceed to investigate whether time-dependent expectation values in truncated methods, given by

$$\langle O(t)\rangle_\parallel = l_\parallel^T(t)O_\parallel(t)r_\parallel(t), \quad (42)$$

have any dependence on the truncated AB partition, as this would imply that the the general scaling properties of the method are incorrect. Assuming that the state of system is known at $t=t_0$, the time-dependent right and transposed left amplitude vectors of truncated methods can be given as exact solutions of the truncated right and left matrix TDSEs in Eq. (14) and Eq. (15),

$$r_\parallel(t) = U_\parallel(t,t_0)r_\parallel(t_0), \quad l_\parallel^T(t) = l_\parallel^T(t_0)U_\parallel(t_0,t), \quad (43)$$

where

$$U_\parallel(t,t_0) = I_\parallel + \\ \sum_{n=0}^\infty (-i)^n \int_{t_0}^t dt_1\int_{t_0}^{t_1}dt_2\cdots\int_{t_0}^{t_{n-1}}dt_n\ \widetilde{H}_\parallel(t_1)\cdots\widetilde{H}_\parallel(t_n) \quad (44)$$

and I is the identity matrix. Inserting Eq. (43) into Eq. (42), we get that

$$\langle O(t)\rangle_\parallel = l_\parallel^T(t_0)U_\parallel(t_0,t)O_\parallel(t)U_\parallel(t,t_0)r_\parallel(t_0). \quad (45)$$

In general, the truncated shifted Hamiltonian matrices taking part in this equation have the same block structure as Eq. (41), namely

$$\widetilde{H}_\parallel(t) = \\ \begin{pmatrix} \widetilde{H}_{00}(t) & \widetilde{H}_{0A}(t) & \widetilde{H}_{0B}(t) & 0 \\ \widetilde{H}_{A0}(t) & \widetilde{H}_{AA}(t) & 0 & (\widetilde{H}_{AAB}(t))_\parallel \\ \widetilde{H}_{B0}(t) & 0 & \widetilde{H}_{BB}(t) & (\widetilde{H}_{BAB}(t))_\parallel \\ 0 & (\widetilde{H}_{ABA}(t))_\parallel & (\widetilde{H}_{ABB}(t))_\parallel & (\widetilde{H}_{ABAB}(t))_\parallel \end{pmatrix}. \quad (46)$$

In TD-EOM-CC, truncated shifted Hamiltonian matrices generally exhibit this block structure. This is because the $U(t)$ term given by Eq. (20) contributes to all upper triangular blocks of Eq. (46) except $\widetilde{H}_{00}(t)$, the $\widetilde{D}(t)$ term given by Eq. (27) to all diagonal blocks of the matrix, and the $\widetilde{L}(t)$ term given by Eq. (28) to all lower triangular blocks of the matrix except $\widetilde{H}_{00}(t)$.

If no further assumptions can be made regarding the block structure of $\widetilde{H}_\parallel(t)$ in Eq. (46), we cannot assume

the correctness of the scaling properties of the method. This becomes apparent when considering that the block matrices in both Eq. (41) and Eq. (46) can map partitions of right and transposed left amplitude vectors to partitions where the numbers of excited subsystems have both increased and decreased by one. Consequently, the products of two or more such matrices participating in $\boldsymbol{U}_\parallel(t_0,t)\boldsymbol{O}_\parallel(t)\boldsymbol{U}_\parallel(t,t_0)$ in Eq. (45) can result in a mapping between the non-zero 0 partition of $\boldsymbol{r}_\parallel(t_0)$ and the truncated AB partition of $\boldsymbol{l}_\parallel^T(t_0)$. This implies that time-dependent expectation values for two non-interacting subsystems, given by Eq. (42), are generally affected by the truncation of the product bases, that is,

$$\langle O(t)\rangle_\parallel \neq \boldsymbol{l}^T(t)\boldsymbol{O}(t)\boldsymbol{r}(t). \tag{47}$$

Thus, the scaling properties of the truncated TD-EOM-CC method are generally incorrect.

In TDCC, the $\widetilde{\boldsymbol{L}}(t)$ term of the shifted Hamiltonian matrix $\widetilde{\boldsymbol{H}}(t)$ is always zero, as seen from Eq. (31). In TD-EOM-CC, the $\widetilde{\boldsymbol{L}}(t)$ term is zero whenever the interaction term $V(t)$ is zero, as seen from Eq. (28). In both these cases, the blocks below the diagonal of $\widetilde{\boldsymbol{H}}_\parallel(t)$ are equal to zero, and Eq. (46) can be written as

$$\widetilde{\boldsymbol{H}}_\parallel(t) = \begin{pmatrix} \widetilde{H}_{00}(t) & \widetilde{\boldsymbol{H}}_{0\,A}(t) & \widetilde{\boldsymbol{H}}_{0\,B}(t) & \boldsymbol{0} \\ \boldsymbol{0} & \widetilde{\boldsymbol{H}}_{A\,A}(t) & \boldsymbol{0} & (\widetilde{\boldsymbol{H}}_{A\,AB}(t))_\parallel \\ \boldsymbol{0} & \boldsymbol{0} & \widetilde{\boldsymbol{H}}_{B\,B}(t) & (\widetilde{\boldsymbol{H}}_{B\,AB}(t))_\parallel \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & (\widetilde{\boldsymbol{H}}_{AB\,AB}(t))_\parallel \end{pmatrix}. \tag{48}$$

The $\widetilde{\boldsymbol{D}}(t)$ term given by Eq. (21) is also zero in TDCC, implying that $\widetilde{H}_{00}(t) = 0$. However, we choose to include the non-zero $\widetilde{H}_{00}(t)$ in the following discussion, where it is seen that this element does not have any impact on the scaling properties of truncated methods. This is to accommodate field-free TD-EOM-CC in the discussion, where $V(t) = 0$ for all times t.

We now examine how the vectors of the initial state are affected by the repeated transformation by shifted Hamiltonian matrices with the block structure of Eq. (48). These matrices do not map partitions of right amplitude vectors to partitions with a higher number of excited subsystems. Consequently, assuming $\boldsymbol{r}(t_0) = \boldsymbol{r}^{(0)} = (1, \boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0})^T$,

$$\widetilde{\boldsymbol{H}}_\parallel(t_1)\cdots\widetilde{\boldsymbol{H}}_\parallel(t_n)\boldsymbol{r}_\parallel(t_0) = \begin{pmatrix} (\widetilde{\boldsymbol{H}}(t_1)\cdots\widetilde{\boldsymbol{H}}(t_n)\boldsymbol{r}(t_0))_0 \\ \boldsymbol{0} \\ \boldsymbol{0} \\ \boldsymbol{0} \end{pmatrix}, \tag{49}$$

where the right-hand side is described purely in terms of untruncated matrices and vectors, which is valid since the AB partition does not contribute to any of the matrix-vector transformations. Furthermore, the same matrices do not map partitions of transposed left amplitude vec-

tors to partitions with a lower number of excited subsystems. Consequently,

$$\boldsymbol{l}_\parallel^T(t_0)\widetilde{\boldsymbol{H}}_\parallel(t_n)\cdots\widetilde{\boldsymbol{H}}_\parallel(t_1) = \begin{pmatrix} (\boldsymbol{l}^T(t_0)\widetilde{\boldsymbol{H}}(t_n)\cdots\widetilde{\boldsymbol{H}}(t_1))_0 \\ (\boldsymbol{l}^T(t_0)\widetilde{\boldsymbol{H}}(t_n)\cdots\widetilde{\boldsymbol{H}}(t_1))_A \\ (\boldsymbol{l}^T(t_0)\widetilde{\boldsymbol{H}}(t_n)\cdots\widetilde{\boldsymbol{H}}(t_1))_B \\ (\boldsymbol{l}_\parallel^T(t_0)\widetilde{\boldsymbol{H}}_\parallel(t_n)\cdots\widetilde{\boldsymbol{H}}_\parallel(t_1))_{AB} \end{pmatrix}^T, \tag{50}$$

where the 0, A and B partitions on the right-hand side are described purely in terms of untruncated matrices and vectors, which is valid since the AB partition does not contribute to any of the corresponding matrix-vector transformations.

Under the condition that the matrix $\widetilde{\boldsymbol{H}}(t)$ has the block upper triangular structure of Eq. (48) for all times t, and that the right amplitude vector starts out as the ground state vector $\boldsymbol{r}(t_0) = \boldsymbol{r}^{(0)}$, Eq. (49) and Eq. (50) imply that

$$\boldsymbol{r}_\parallel(t) = \begin{pmatrix} (\boldsymbol{U}(t,t_0)\boldsymbol{r}(t_0))_0 \\ \boldsymbol{0} \\ \boldsymbol{0} \\ \boldsymbol{0} \end{pmatrix}, \tag{51}$$

$$\boldsymbol{l}_\parallel^T(t) = \begin{pmatrix} (\boldsymbol{l}^T(t_0)\boldsymbol{U}(t_0,t))_0 \\ (\boldsymbol{l}^T(t_0)\boldsymbol{U}(t_0,t))_A \\ (\boldsymbol{l}^T(t_0)\boldsymbol{U}(t_0,t))_B \\ (\boldsymbol{l}_\parallel^T(t_0)\boldsymbol{U}_\parallel(t_0,t))_{AB} \end{pmatrix}^T, \tag{52}$$

where $\boldsymbol{U}(t,t_0)$ is the untruncated counterpart of the time evolution operator in Eq. (44), containing the untruncated time-dependent Hamiltonian matrices $\widetilde{\boldsymbol{H}}(t)$ evaluated at different times t.

To assess the effect of truncation on time-dependent expectation values in TDCC and field-free TD-EOM-CC, we note that the matrix $\boldsymbol{O}_\parallel(t)$ does not map between the 0 partition of right amplitude vectors and the AB partition of transposed left amplitude vectors. Hence, time-dependent expectation values for systems with two non-interacting subsystems, given by Eq. (45), are unaffected by the truncation of the product bases,

$$\begin{aligned} \langle O(t)\rangle_\parallel &= \boldsymbol{l}_\parallel^T(t_0)\boldsymbol{U}_\parallel(t_0,t)\boldsymbol{O}_\parallel(t)\boldsymbol{U}_\parallel(t,t_0)\boldsymbol{r}_\parallel(t_0) \\ &= (\boldsymbol{l}^T(t_0)\boldsymbol{U}(t_0,t))_0 O_{0\,0}(t)(\boldsymbol{U}(t,t_0)\boldsymbol{r}(t_0))_0 \\ &\quad + (\boldsymbol{l}^T(t_0)\boldsymbol{U}(t_0,t))_A \boldsymbol{O}_{A\,0}(t)(\boldsymbol{U}(t,t_0)\boldsymbol{r}(t_0))_0 \\ &\quad + (\boldsymbol{l}^T(t_0)\boldsymbol{U}(t_0,t))_B \boldsymbol{O}_{B\,0}(t)(\boldsymbol{U}(t,t_0)\boldsymbol{r}(t_0))_0 \\ &= \boldsymbol{l}^T(t)\boldsymbol{O}(t)\boldsymbol{r}(t). \end{aligned} \tag{53}$$

Furthermore, time-dependent expectation values can be shown to be unaffected by the truncation of the product bases for any number of non-interacting subsystems, by sequentially dividing one of the remaining composite subsystems into two subsystems and reapplying the

aforementioned arguments. This implies that truncated TDCC and truncated field-free TD-EOM-CC methods have the correct scaling properties when the system starts out in the ground state.

III.   COMPUTATIONAL DETAILS

In order to investigate the properties of the methods described in Section II B numerically, we truncate the methods at the CCSD level, giving the TD-EOM-CCSD and TDCCSD methods. The TD-EOM-CCSD method is implemented in the spin-adapted elementary basis in a development version of the e^T program [27]. Furthermore, we use the existing implementations of the spin-adapted ground state and TDCCSD methods in e^T 1.0 [27, 28]. The methods are used to calculate the interaction of atoms with the electromagnetic field represented by the electric field

$$\boldsymbol{\mathcal{E}}(t) = \mathcal{E}_0 \boldsymbol{\epsilon} \cos(\omega_0(t - t_0) + \phi) f(t) \qquad (54)$$

where ϵ_0 is the peak field strength, $\boldsymbol{\epsilon}$ the polarization, ω_0 the carrier frequency, ϕ the carrier-envelope phase and $f(t)$ the envelope of the field. The envelope is given the functional form

$$f(t) = \begin{cases} 0, & t < a, \\ \sin^2\left(\frac{2\pi(t-a)}{4(b-a)}\right), & a \le t \le b, \\ 1, & t > b, \end{cases} \qquad (55)$$

which increases from zero to one in the interval from a to b.

The aug-cc-pVDZ basis set is used for the helium and beryllium atoms in the simulations. The field is given a carrier frequency of $1.005\,749\,62$ a.u., which corresponds to the transition between the ground $0\,^1S_0$ state and the first dipole-allowed excited $2\,^1P_1$ state of helium calculated with EOM-CCSD and aug-cc-pVDZ. By adopting this carrier frequency, the field is in resonance with the same transition for a single helium atom in both the TD-EOM-CCSD and TDCCSD simulations. The field is furthermore given a peak field strength of 3×10^{-2} a.u., a polarization in the z-direction and a carrier-envelope phase of $\phi = -\pi/2$. The envelope of the field is set to increase from $a = 0$ a.u. of time until $b = 15$ optical cycles (≈ 93.709 a.u.). The envelope gives the field a narrow bandwidth, centered around the $0\,^1S_0$–$2\,^1P_1$ resonance, which ensures that the time-dependent state of single-helium simulations is kept in a superposition dominated by the two states. The integration of the time-dependent differential equations is done using the Dormand-Prince method of order 5(4) with the adaptive time stepping procedure described in Appendix B of Ref. [24]. The initial time step size is set to 1×10^{-2} a.u., and the maximum and minimum values of the estimated error are set to 1×10^{-9} a.u. and 1×10^{-11} a.u., respectively. The integration is stopped when the size of the adaptive time step becomes smaller than 1×10^{-10}.
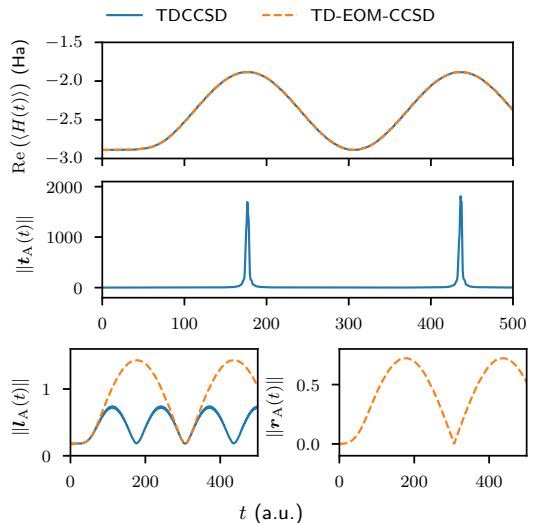

FIG. 1. Real part of the time-dependent energy expectation value, Re$\left(\langle H(t)\rangle\right)$, and the right and left 2-norms of the A partitions of the cluster, right and left amplitudes, $\|\boldsymbol{t}_A(t)\|$, $\|\boldsymbol{r}_A(t)\|$ and $\|\boldsymbol{l}_A(t)\|$, from time-dependent TD-EOM-CCSD and TDCCSD simulations of a single helium atom in an external field.

IV.   RESULTS AND DISCUSSION

A.   Simulating single-subsystem Rabi oscillations with TD-EOM-CCSD and TDCCSD

For a single helium atom, both the TD-EOM-CCSD and TDCCSD methods can describe all possible excitations of the reference determinant, and the time-dependent observables are thus analytically equal for the two methods. In the top panel of Fig. 1, we show that this is also the case numerically for the time-dependent energy expectation value, as the two methods give the same result. The expectation value is initially equal to the ground state energy, and periodically increases and decreases as a function of time, illustrating that the system undergoes Rabi oscillation between the $0\,^1S_0$ and $2\,^1P_1$ states of the helium atom. The time-dependent 2-norms of the cluster, right and left amplitudes are also shown in Fig. 1. In contrast to the right and left amplitude norms of both methods, the norm of the A partition of the cluster amplitudes in TDCCSD displays narrow peaks which coincide with the maxima of the time-dependent energy expectation value. TDCCSD is known to be numerically unstable when the weight of the reference determinant approaches zero [20, 21], but we observe that the method can still be used for simulating Rabi oscillations between the $0\,^1S_0$ and $2\,^1P_1$ states of a single helium atom.
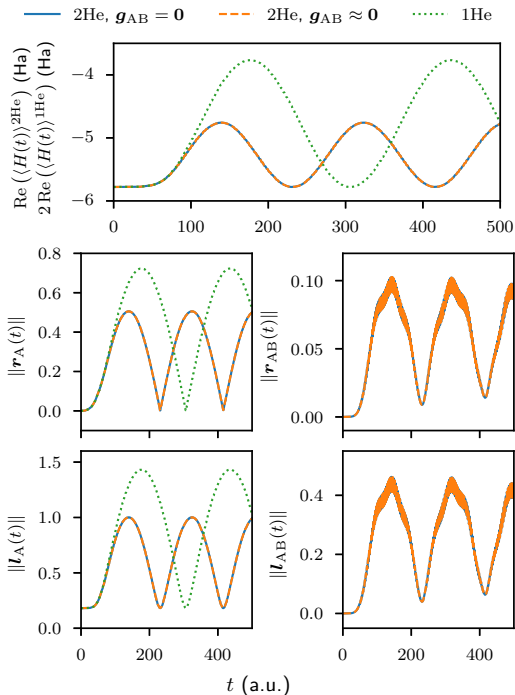
FIG. 2. TD-EOM-CCSD simulations of one (1He) and two distant (2He) helium atoms in an external field. Results from two different two-helium simulations are shown, one performed with the regular TD-EOM-CCSD method ($g_{AB} \approx 0$), and one where the two-center Coulomb integrals have been set to zero ($g_{AB} = 0$). The single-helium results in Fig. 1 are also shown for comparison. In the top panel, the real part of the time-dependent energy expectation value, Re $\left(\langle H(t) \rangle \right)$, is shown, where the single-helium result has been multiplied by two to better assess the scaling properties of the method. In the first column below the top panel, the 2-norms of the A partitions of the right and left amplitudes, $\|\boldsymbol{r}_A(t)\|$ and $\|\boldsymbol{l}_A(t)\|$, are shown. In the next column, the 2-norms of the AB partitions of the right and left amplitudes, $\|\boldsymbol{r}_{AB}(t)\|$ and $\|\boldsymbol{l}_{AB}(t)\|$, are shown.

B. Simulating two-subsystem Rabi oscillations with TD-EOM-CCSD

To numerically investigate the scaling properties of TD-EOM-CCSD, we simulate the interaction between two helium atoms and the external field, where one of the atoms is located at the origin and the other at $1 \times 10^7 \, a_0$ along the x-axis. The large interatomic separation is chosen in order to minimize any interaction between the atoms, without requiring any modifications to the integral code. At this separation, the only non-zero two-center electronic integrals are the two-center Coulomb integrals, since the other integrals rapidly decay below

the default integral cutoffs of e^T due to their dependence on the overlap of the orbitals on the different centers. The two-center Coulomb integrals decay as the inverse distance between the electrons on the two centers, and have magnitudes of less than $1 \times 10^{-6} \, E_h$ in the simulations, implying that $H(t) \approx H_A(t) \otimes I_B + I_A \otimes H_B(t)$.

The simulation successfully completes, and the time-dependent energy expectation value and various time-dependent amplitude 2-norms are shown together with the single-helium results of Section IV A in Fig. 2. The single-helium energy expectation value has been multiplied by a factor of two, since the correct scaling properties would require the two-helium energy expectation value to be equal to two times the single-helium one. The frequency of the oscillations in the scaled energy expectation value and single-subsystem norms increases, and their magnitude decreases, as the number of helium atoms increases from one to two.

A separate TD-EOM-CCSD simulation is performed where the two-center Coulomb integrals are eliminated. The results are shown in Fig. 2, demonstrating that the elimination of the integrals does not have any significant impact on the results. It further demonstrates that the incorrect scaling properties of TD-EOM-CCSD is not due to the weak interaction between the distant atoms, but rather a problem caused the insufficient flexibility of the parametrization of the method.

C. Simulating two-subsystem Rabi oscillations with TDCCSD

To numerically investigate the scaling properties of TDCCSD, we start by comparing the results from the single-helium simulation in Section IV A with results from simulations of two helium atoms, where one of the atoms is located at the origin and the other at $1 \times 10^7 \, a_0$ along the x-axis. The large interatomic separation implies that Eq. (39) is approximately satisfied at the start of the simulation.

In Fig. 3, the time-dependent energy expectation value and 2-norms of various amplitude partitions are shown, for different two-helium TDCCSD simulations, and compared to the results from the single-helium calculation in Section IV A. From the start of the simulation and up to 170 a.u. of time, the regular TDCCSD method gives an energy expectation value that is equal to two times the expectation value from the single-helium simulation, since the absolute difference between these quantities are on the order of 1×10^{-14} and less. This agrees with the theory in Section II C, and implies that the TDCCSD method treats the correlation exactly in this time range, even though the system has four electrons and the operators in Eq. (2) and Eq. (3) are truncated at the doubles level. The two-helium single-subsystem norms are also essentially equal to the single-helium norms in the same interval, as illustrated by the small absolute differences between these quantities. After around 170 a.u.
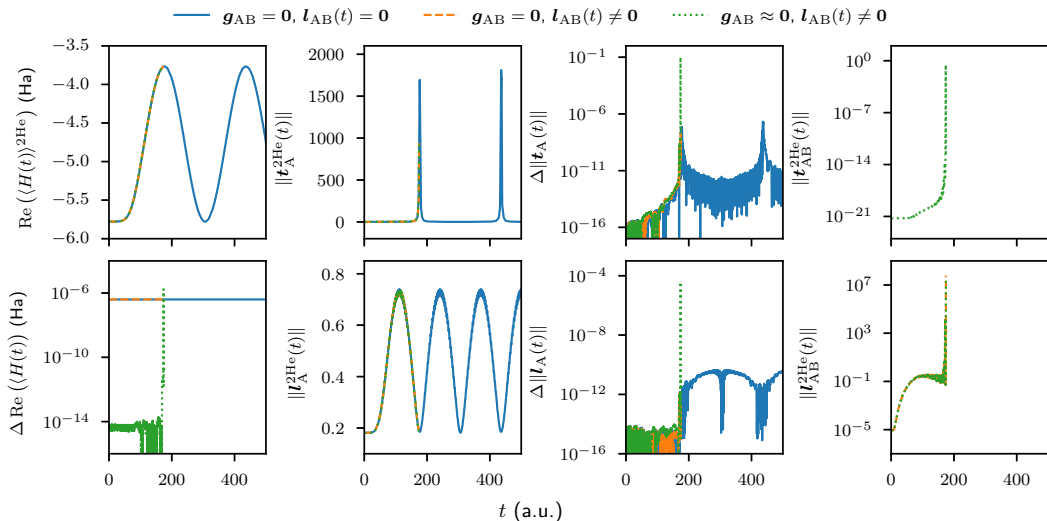
FIG. 3. TDCCSD simulations of one (1He) and two distant (2He) helium atoms in an external field. Results from three different two-helium simulations are shown, one performed with the regular TDCCSD method ($g_{AB} \approx 0$, $l_{AB}(t) \neq 0$), one where the two-center Coulomb integrals have been set to zero ($g_{AB} = 0, l_{AB}(t) \neq 0$), and one where both the two-center Coulomb integrals and the AB partition of the left amplitudes have been set to zero ($g_{AB} = 0, l_{AB}(t) = 0$). The first column of panels shows the real part of the time-dependent energy expectation value $\text{Re}\left(\langle H(t)\rangle^{2He}\right)$ and the absolute difference $\Delta\,\text{Re}\left(\langle H(t)\rangle\right) = \left|\text{Re}\left(\langle H(t)\rangle^{2He}\right) - 2\,\text{Re}\left(\langle H(t)\rangle^{1He}\right)\right|$ with the single-helium result multiplied by a factor of two. In the second column, the 2-norms of the A partitions of the cluster and left amplitudes $\|t_A^{2He}(t)\|$ and $\|l_A^{2He}(t)\|$ are shown, and in the third column the absolute differences $\Delta\|t_A(t)\| = \left|\|t_A^{2He}(t)\| - \|t_A^{1He}(t)\|\right|$ and $\Delta\|l_A(t)\| = \left|\|l_A^{2He}(t)\| - \|l_A^{1He}(t)\|\right|$ with the single-helium norms. In the last column, the 2-norms of the AB partitions of the cluster and left amplitudes, $\|t_{AB}^{2He}(t)\|$ and $\|l_{AB}^{2He}(t)\|$, are shown.

of time, however, the regular TDCCSD simulation fails, and the integration stops as adaptive time steps smaller than 1×10^{-10} are needed to proceed. Toward the end of the simulation, the absolute differences between the two-helium and scaled single-helium results blow up by increasing several orders of magnitude, implying that the simulation does not behave according to the description in Section II C. Moreover, the norms of the AB partitions of both the cluster and left amplitudes rise throughout the simulation, and also increase dramatically at the end. In the following, we argue that the blowup of the AB partitions of the amplitudes are the reason for the failure of the regular TDCCSD simulation.

The observed increase in the norm of the AB partition of the cluster amplitudes indicates that the conditions for Eq. (39) are not met exactly. To enforce this equation, a separate TDCCSD simulation is conducted where all two-center Coulomb integrals are eliminated, which indeed results in the norm for the AB partition of the cluster amplitudes being equal to zero for all times t. However, the simulation still fails after around 170 a.u., and the 2-norm of the AB partition of the left amplitudes AB increases dramatically toward the end of the

simulation, as shown in Fig. 3. The energy expectation value and single-subsystem 2-norms are also shown and compared to the scaled single-helium results in the figure. In contrast to the regular TDCCSD simulation, there are no visible instabilities in the absolute differences between the two-helium and scaled single-helium results, although the absolute difference in the energy expectation value has increased to a constant on the order of 1×10^{-6} E_h. This difference is larger than the one seen in the initial stages of the regular TDCCSD simulation, and can be explained by the disruption of the equilibrium between the attractive and repulsive forces between the two atoms caused by the omission of the two-center Coulomb integrals. The difference can be expected to be smaller if all two-center interactions are removed from the integral code. Overall, the results support the correctness of the scaling properties of the TDCC method in the limit of no subsystem interaction, but also highlight the sensitivity of the cluster amplitudes to interactions between the two subsystems in simulations of collective Rabi oscillations.

Provided the AB partition of the cluster amplitudes is zero in TDCC, Eq. (53) predicts that time-dependent expectation values should not depend on the AB partition
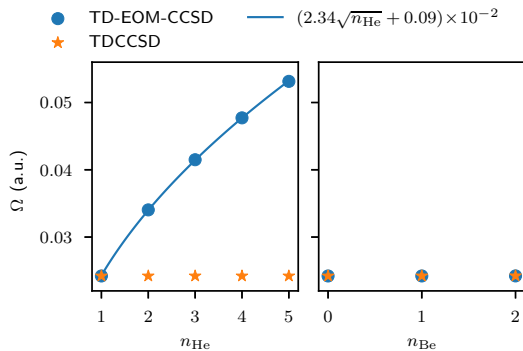
FIG. 4. TD-EOM-CCSD and TDCCSD simulations of distant helium and beryllium atoms in an external field, with the two-center Coulomb integrals set to zero. The left panel shows results for one to five helium atoms, and the right panel results for one helium atom and zero to two beryllium atoms. The Rabi frequencies Ω are obtained by least-squares fitting the function $A\sin(\Omega t + \varphi) + C$ to the real part of the time-dependent energy expectation value. The function $(2.34\sqrt{n_{\text{He}}} + 0.09) \times 10^{-2}$ is obtained by least-squares fitting the function $A\sqrt{n_{\text{He}}} + C$ to the TD-EOM-CCSD Rabi frequencies in the left panel.

of the left amplitudes. To verify this prediction and evaluate the impact of eliminating the AB partition of both vectors on the numerical stability, a separate simulation is conducted where the two-center Coulomb integrals and the initial values and time derivatives of the AB partition of the left amplitudes are set to zero. This simulation completes successfully, demonstrating that the failure of the previous simulation is caused by the AB partition of the left amplitudes. In Fig. 3, the resulting energy expectation value and single-subsystem 2-norms are shown and compared to the scaled single-helium results. The absolute differences between the two-helium and scaled single-helium results are consistently below 1×10^{-6}, and there is no visible difference in the energy expectation value caused by the removal of the AB partition of the left amplitudes, in agreement with the theoretical observations made in Section II C.

D. Scaling of collective Rabi oscillations with the number of subsystems

To further investigate the scaling properties of TD-EOM-CCSD and TDCCSD, the interaction with the field is also calculated for three to five helium atoms, and for one helium atom together with one to two beryllium atoms. The first helium atom is located at the origin, and the following atoms at increments of $1 \times 10^7\, a_0$ along the x-axis. In all simulations, the two-center Coulomb integrals are eliminated to circumvent the blowup of the AB partition of the cluster amplitudes in TDCCSD and

treat the two methods on an equal footing. Furthermore, the initial values and time derivatives of the AB partition of the left amplitudes is also set to zero in all TDCCSD simulations. All simulations complete successfully.

To estimate the Rabi frequencies Ω from the simulation results, the sinusoidal function $A\sin(\Omega t + \varphi) + C$ is least-squares fitted to the real part of the oscillating time-dependent energy expectation value between $t = 100$ a.u. and $t = 500$ a.u.. The resulting frequencies are displayed in Fig. 4. From the figure, we can see that the Rabi frequencies from the TD-EOM-CCSD simulations increase with the number of helium atoms. The function $A\sqrt{n_{\text{He}}} + C$ is least-squares fitted to these frequencies to evaluate their scaling properties, where n_{He} represents the number of helium atoms. The fitted curve, also illustrated in Fig. 4, demonstrates that the Rabi frequency increases as the square root of the total number of helium atoms. As the number of subsystems in resonance with the field increases, the frequency can therefore erroneously appear to approach infinity, meaning that the TD-EOM-CCSD method gives a qualitatively incorrect representation of transitions occurring in multiple subsystems simultaneously. This scaling behavior is similar to the physical square-root scaling that can be observed when a single excitation is symmetrically shared among N atoms experiencing effectively uniform coupling and interaction, such as when multi-atom Rydberg excitations are blocked by van der Waals interactions [29]. In our simulations, however, the scaling effect is caused by the limited flexibility of the TD-EOM-CCSD parametrization, as removing the two-center Coulomb integrals has no discernible impact on the results in Section IV B.

For all TDCCSD simulations, the Rabi frequencies shown in Fig. 4 remain constant regardless of the number of helium or beryllium atoms, as predicted in Section II C. In addition, the Rabi frequencies also remain constant in TD-EOM-CCSD simulations involving a helium atom and up to two beryllium atoms, suggesting that Rabi frequencies in TD-EOM-CC are insensitive to the number of off-resonant subsystems. Therefore, TD-EOM-CCSD may provide an accurate portrayal of a solitary Rabi oscillation in an extended quantum system while maintaining superior numerical stability compared to TDCCSD.

V. CONCLUSION

In this work, a general theoretical framework for representing both the TD-EOM-CC and TDCC methods was presented, incorporating the time derivatives of the cluster amplitudes as auxiliary conditions. Through this framework, it was demonstrated that the scaling properties of truncated TD-EOM-CC methods are incorrect in general, while the scaling properties are correct for truncated TDCC and field-free TD-EOM-CC methods.

The TD-EOM-CCSD method was implemented in the elementary basis, and used to numerically compare the

scaling properties of the TD-EOM-CCSD and TDCCSD methods through simulations of collective Rabi oscillations. The simulations revealed that the TD-EOM-CCSD method gives a qualitatively incorrect representation of collective Rabi oscillations, since the Rabi frequency increases with the number of subsystems in resonance with the external field. Nevertheless, the results also indicate that truncated TD-EOM-CC methods can be suitable for simulating solitary Rabi oscillations in extended quantum systems.

All TD-EOM-CCSD simulations completed successfully, while the regular TDCC simulations of collective Rabi oscillations in distant subsystems failed due to the cluster and left amplitudes blowing up. Despite this, the initial stages of the TDCCSD simulations displayed the correct scaling properties, as predicted by Section II C, suggesting that TDCC can be suitable for simulating collective transitions in extended systems, as long as complete population inversion is avoided. Furthermore, it was demonstrated that the numerical stability of the TDCC simulations can be enhanced by ensuring that all two-center integrals are precisely zero and eliminating the two-subsystem deexcited partition of the left amplitudes. However, this approach is not suitable when simulating collective Rabi oscillations in systems that necessitate the description of interaction between subsystems.

In conclusion, we propose that further research should be dedicated to the development of approximate methods that can provide a qualitatively correct description of collective Rabi oscillations, even when the interaction between the subsystems is non-zero.

ACKNOWLEDGMENTS

[1] T. Brabec and F. Krausz, Rev. Mod. Phys. **72**, 545 (2000).

[2] E. A. Seddon, J. A. Clarke, D. J. Dunning, C. Masciovecchio, C. J. Milne, F. Parmigiani, D. Rugg, J. C. H. Spence, N. R. Thompson, K. Ueda, S. M. Vinko, J. S. Wark, and W. Wurth, Rep. Prog. in Phys. **80**, 115901 (2017).

[3] J. Duris, S. Li, T. Driver, E. G. Champenois, J. P. MacArthur, A. A. Lutman, Z. Zhang, P. Rosenberger, J. W. Aldrich, R. Coffee, G. Coslovich, F.-J. Decker, J. M. Glownia, G. Hartmann, W. Helml, A. Kamalov, J. Knurr, J. Krzywinski, M.-F. Lin, J. P. Marangos, M. Nantel, A. Natan, J. T. O'Neal, N. Shivaram, P. Walter, A. L. Wang, J. J. Welch, T. J. A. Wolf, J. Z. Xu, M. F. Kling, P. H. Bucksbaum, A. Zholents, Z. Huang, J. P. Cryan, and A. Marinelli, Nat. Photon **14**, 30 (2020).

[4] V. Stooß, S. M. Cavaletto, S. Donsa, A. Blättermann, P. Birk, C. H. Keitel, I. Březinová, J. Burgdörfer, C. Ott, and T. Pfeifer, Phys. Rev. Lett. **121**, 173005 (2018).

[5] Y. Silberberg, Annu. Rev. Phys. Chem. **60**, 277 (2009).

[6] K. Kulander and M. Lewenstein, Multiphoton and strong-field processes, in *Springer Handbook of Atomic, Molecular, and Optical Physics*, edited by G. Drake (Springer New York, New York, NY, 2006) pp. 1077–1089.

[7] T. Bayer, M. Wollenhaupt, and T. Baumert, J. Phys. B: Atom. Mol. Phys. **41**, 074007 (2008).

[8] A. Hishikawa, A. Matsuda, and M. Fushitani, Bull. Chem. Soc. Jpn. **93**, 1293 (2020).

[9] I. R. Solá, J. González-Vázquez, R. de Nalda, and L. Bañares, Phys. Chem. Chem. Phys. **17**, 13183 (2015).

[10] A. Palacios and F. Martín, WIREs Comput. Mol. Sci. **10**, e1430 (2020).

[11] H. R. Larsson, S. Bauch, L. K. Sørensen, and M. Bonitz, Phys. Rev. A **93**, 013426 (2016).

[12] M. Labeye, F. Zapata, E. Coccia, V. Véniard, J. Toulouse, J. Caillat, R. Taïeb, and E. Luppi, J. Chem. Theory Comput. **14**, 5846 (2018).

[13] X. Li, N. Govind, C. Isborn, A. E. DePrince, and K. Lopata, Chem. Rev. **120**, 9951 (2020).

[14] R. J. Bartlett, in *Theory and Applications of Computational Chemistry*, edited by C. E. Dykstra, G. Frenking, K. S. Kim, and G. E. Scuseria (Elsevier, Amsterdam, 2005) pp. 1191–1221.

[15] F. A. Evangelista, J. Chem. Phys. **149**, 030901 (2018).

[16] R. J. Bartlett, Annu. Rev. Phys. Chem. **32**, 359 (1981).

[17] H. Koch, H. J. A. Jensen, P. Jørgensen, and T. Helgaker, J. Chem. Phys. **93**, 3345 (1990).

[18] H. Koch, R. Kobayashi, A. Sanchez de Merás, and P. Jørgensen, J. Chem. Phys. **100**, 4393 (1994).

[19] M. Nooijen, K. R. Shamasundar, and D. Mukherjee, Mol. Phys. **103**, 2277 (2005).

[20] T. B. Pedersen and S. Kvaal, J. Chem. Phys. **150**, 144106 (2019).

[21] H. E. Kristiansen, Ø. S. Schøyen, S. Kvaal, and T. B. Pedersen, J. Chem. Phys. **152**, 071102 (2020).

[22] J. A. Sonk, M. Caricato, and H. B. Schlegel, J. Phys. Chem. A **115**, 4678 (2011).

[23] E. Luppi and M. Head-Gordon, Mol. Phys. **110**, 909 (2012).

[24] A. S. Skeidsvoll, T. Moitra, A. Balbi, A. C. Paul, S. Coriani, and H. Koch, Phys. Rev. A **105**, 023103 (2022).

[25] H. Koch and P. Jørgensen, J. Chem. Phys. **93**, 3333 (1990).

[26] T. B. Pedersen, H. E. Kristiansen, T. Bodenstein, S. Kvaal, and Ø. S. Schøyen, J. Chem. Theory Comput. **17**, 388 (2021).

[27] S. D. Folkestad, E. F. Kjønstad, R. H. Myhre, J. H. Andersen, A. Balbi, S. Coriani, T. Giovannini, L. Goletto, T. S. Haugland, A. Hutcheson, I.-M. Høyvik, T. Moitra, A. C. Paul, M. Scavino, A. S. Skeidsvoll, Å. H. Tveten, and H. Koch, J. Chem. Phys. **152**, 184103 (2020).

[28] A. S. Skeidsvoll, A. Balbi, and H. Koch, Phys. Rev. A **102**, 023115 (2020).

[29] J. Zeiher, P. Schauß, S. Hild, T. Macrì, I. Bloch, and C. Gross, Phys. Rev. X **5**, 031015 (2015).

Paper V

Coupled cluster simulation of impulsive stimulated x-ray Raman scattering

Alice Balbi,[1, *] Andreas S. Skeidsvoll,[2, *] and Henrik Koch[1, 2, †]

[1]*Scuola Normale Superiore, Piazza dei Cavalieri, 7, I-56126, Pisa, Italy*
[2]*Department of Chemistry, Norwegian University of Science and Technology, 7491 Trondheim, Norway*
(Dated: May 31, 2023)

Time-dependent equation-of-motion coupled cluster (TD-EOM-CC) is used to simulate impulsive stimulated x-ray Raman scattering (ISXRS) of ultrashort laser pulses by neon, carbon monoxide, pyrrole, and p-aminophenol. The TD-EOM-CC equations are expressed in the basis of field-free EOM-CC states, where the calculation of the core-excited states is simplified through the use of the core-valence separation (CVS) approximation. The transfer of electronic population from the ground state to the core- and valence-excited states is calculated for different numbers of included core- and valence-excited states, as well as for electric field pulses with different polarizations and carrier frequencies. The results indicate that Gaussian pulses can transfer significant electronic populations to the valence states through the Raman process. The sensitivity of this population transfer to the model parameters is analyzed. The time-dependent electronic density for p-aminophenol is also showcased, supporting the interpretation that ISXRS involves localized core excitations and can be used to rapidly generate valence wavepackets.

I. INTRODUCTION

The ability to experimentally generate short and intense x-ray laser pulses has been a subject of significant interest in the field of x-ray science. Recent technological advances, specifically the realization of x-ray free electron lasers (XFELs) [1, 2] and new approaches based on high harmonic generation (HHG) [3, 4], have made it possible to generate x-ray laser pulses with high intensities and pulse durations as short as a few hundred and even tens of attoseconds [5]. This progress has enabled the development of new experimental techniques with unprecedented temporal resolution, facilitating the imaging and control of atoms and molecules on the time scale of electronic motion. [6–12] An important phenomenon in this context is impulsive stimulated x-ray Raman scattering (ISXRS), which is the extension of stimulated x-ray Raman scattering (SXRS) to the impulsive limit, where the duration of the external field interaction is short compared to the time scales of the subsequent evolution of the system.

In general, Raman scattering is a light-matter interaction phenomenon in which photons trigger an excitation of an atomic or molecular system followed by a deexcitation to an energy level different from the initial one. In the context of x-ray Raman scattering, the involved transitions are electronic in character. [6, 13–16] We focus on the situation in which the electronic excitation in play is a core excitation, which is deexcited to a valence-excited state through the decay of a valence electron into a core vacancy, see Fig. 1. Core excitations are often localized on a specific atomic site and sensitive to the surrounding electronic environment, making them useful for the local initiation of charge migration. We treat the case where both the excitation and deexcitation are stimulated by an interaction with the same laser pulse. [17]
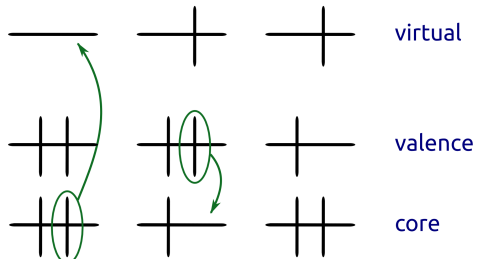

FIG. 1. Illustration of the steps in the ISXRS process. Initially, the molecule is in its ground state (left). An external x-ray pulse excites a core electron, leading to a core-excited state (middle). The same pulse can trigger the decay of a valence electron into the core vacancy, leading to a valence-excited state (right).

This is achievable by utilizing a pulse with sufficient bandwidth to encompass the energy differences between the ground state and the core-excited states of interest, as well as between these core-excited states and the final valence-excited states. The interaction with such pulses is similar to the interactions occurring in the first experimental demonstration of electronic population transfer via ISXRS, which was made for the NO molecule at the Linac Coherent Light Source as recently as in 2020. [18]

The progress in experimental techniques has stimulated the development of methods for modeling electron dynamics based on the time-dependent Schrödinger equation. Real-time methods, which involve solving this equation in the real time domain, offer a particularly suitable approach for analyzing ultrafast phenomena. [19] Among these methods, real-time coupled coupled cluster methods offer high accuracy and computational costs that scale polynomially with the system size.

The time evolution is described by differential equations that can be solved using standard numerical integration techniques such as Runge-Kutta methods.

A specific subcategory of real-time coupled cluster methods is the time-dependent coupled cluster (TDCC) methods, [20–28] where the time dependence is parametrized by cluster amplitudes and Lagrange multipliers. [29, 30] These methods offer the advantage of size-extensivity at all levels of truncation. Another subcategory, the time-dependent equation-of-motion coupled cluster (TD-EOM-CC) methods, [31–36] provides less potential for numerical issues compared to TDCC methods, [37] since the time dependence is parametrized by the linear coefficients used in EOM-CC methods and the cluster amplitudes remain fixed at their time-independent ground state values. [38–40]

In the basis of field-free EOM-CC states, the TD-EOM-CC method requires the predetermination of the excited states that are involved in the studied processes. Computationally, the exterior eigenvalue algorithms usually employed for calculating valence-excited states are inefficient for the calculation of the core-excited states often involved in x-ray interactions. This is because the core-excited states have large eigenvalues, and the states are embedded in an ionization (pseudo-)continuum. [41] A useful scheme for the study of core excitations is the core valence separation (CVS) scheme, which disregards all excitations that do not involve at least one core orbital. [42, 43] This allows for the approximate core-excited states to be calculated as the lowest energy states within the reduced excitation space.

In this article, we use the TD-EOM-CC method together with the CVS approximation to simulate the interaction of neon, carbon monoxide, pyrrole, and p-aminophenol with ultrashort laser pulses, and calculate the populations of the valence-excited states following ISXRS targeting molecular K-edges. The article is organized as follows. In Section II we briefly outline the theory behind the calculations. We provide details of the performed computations in Section III, and present and discuss the results in Section IV. Conclusions are presented in Section V.

II. THEORY

The time-dependent system is described by the Hamiltonian

$$H(t) = H^{(0)} + V(t), \qquad (1)$$

where $H^{(0)}$ is the electronic Hamiltonian of the molecule in the Born-Oppenheimer approximation. We describe the interaction with the external laser field $V(t)$ in the dipole approximation and length gauge,

$$V(t) = -\boldsymbol{d} \cdot \boldsymbol{\mathcal{E}}(t), \qquad (2)$$

where $\boldsymbol{d}$ is the vector of Cartesian dipole operators, and $\boldsymbol{\mathcal{E}}(t)$ the Cartesian electric field vector.

The eigenstates of the field-free Hamiltonian,

$$|\psi_j\rangle = \sum_\lambda e^T |\lambda\rangle \, r_{\lambda j} \qquad (3)$$

$$\langle\psi_i| = \sum_\kappa l_{i\kappa} \langle\kappa| \, e^{-T} \qquad (4)$$

can be found by first solving the ground state coupled cluster equations

$$\langle\mu| e^{-T} H^{(0)} e^T |\mathrm{HF}\rangle = 0, \qquad (5)$$

which determine the cluster amplitudes t_μ in the cluster operator,

$$T = \sum_\mu t_\mu \tau_\mu. \qquad (6)$$

Thereafter, the right and left vectors can be found as eigenvectors of the projected time-independent Schrödinger equation,

$$\sum_\lambda \langle\kappa| e^{-T} H^{(0)} e^T |\lambda\rangle \, r_{\lambda j} = r_{\kappa j} E_j, \qquad (7)$$

$$\sum_\kappa l_{i\kappa} \langle\kappa| e^{-T} H^{(0)} e^T |\lambda\rangle = E_i l_{i\lambda}. \qquad (8)$$

These equations lead to the following eigenvalue problems [44]

$$\mathbf{A R}_j = \mathbf{R}_j \Delta E_j, \qquad (9)$$

$$\mathbf{L}_i^T \mathbf{A} = \Delta E_i \mathbf{L}_i^T, \qquad (10)$$

where $A_{\mu\nu} = \langle\mu| e^{-T} \big[H^{(0)}, \tau_\nu\big] e^T |\mathrm{HF}\rangle$, $L_{i\mu} = l_{i\mu}$ and $R_{\nu j} = r_{\nu j}$ for $\mu > 0$ and $\nu > 0$. The excitation energy $\Delta E_j = E_j - E_0$ is given as the difference between the excited state energy and the ground state energy

$$E_0 = \langle\mathrm{HF}| e^{-T} H e^T |\mathrm{HF}\rangle. \qquad (11)$$

The TD-EOM-CC ket and bra states can be expanded in the field-free EOM-CC kets and bras, $|\Psi(t)\rangle = \sum_j |\psi_j\rangle \, c_j(t)$ and $\langle\widetilde{\Psi}(t)| = \sum_i b_i(t) \langle\widetilde{\psi}_i|$. This gives the TD-EOM-CC equations [45]

$$i\frac{\mathrm{d}c_i(t)}{\mathrm{d}t} = \sum_j H_{ij}(t)c_j(t), \qquad (12)$$

$$-i\frac{\mathrm{d}b_j(t)}{\mathrm{d}t} = \sum_i b_i(t)H_{ij}(t), \qquad (13)$$

where $H_{ij}(t) = \langle\widetilde{\psi}_i| H(t) |\psi_j\rangle = \delta_{ij}E_j + \langle\widetilde{\psi}_i| V(t) |\psi_j\rangle$. The time-dependent population of EOM-CC state i in the TD-EOM-CC superposition state can be found as the product of the projections onto the ket and bra of the EOM-CC state,

$$\begin{aligned} P_i(t) &= \langle\widetilde{\Psi}(t)|\psi_i\rangle \, \langle\widetilde{\psi}_i|\Psi(t)\rangle \\ &= b_i(t)c_i(t). \end{aligned} \qquad (14)$$

The eigenvalues of core-excited states are interior to the spectrum of the molecular Hamiltonian, and often hard to reach using exterior eigenvalue methods like Davidson or Lanczos algorithms. The core-valence separation (CVS) approximation [42, 46] simplifies the calculation of these states by removing the valence-core and core-valence blocks of the Hamiltonian and has become a vital tool for the calculation of NEXAFS spectra. [41] Let I denote the set indexing the core orbitals. We invoke the CVS approximation through a projector $\mathcal{P}_I^{\mathrm{CVS}}$ that removes all vector elements that do not reference excitations from at least one core orbital, in each eigensolver iteration. [43] For the coupled cluster singles and doubles (CCSD) truncation level, this can be expressed in compact form as

$$\mathcal{P}_I^{\mathrm{CVS}} r_i^a = l_i^a \mathcal{P}_I^{\mathrm{CVS}} = 0, \quad i \notin I \tag{15}$$

$$\mathcal{P}_I^{\mathrm{CVS}} r_{ij}^{ab} = l_{ij}^{ab} \mathcal{P}_I^{\mathrm{CVS}} = 0, \quad i \notin I \land j \notin I. \tag{16}$$

This projection is effectively setting all elements of the valence-valence block of the full-space elementary basis EOM-CC Jacobian matrix $\mathbf{A}$ to zero, giving the CVS approximated Jacobian matrix, $\mathbf{A}^{\mathrm{CVS}}$. The core-excited EOM-CC states obtained in the CVS approximation can have a non-zero overlap with EOM-CC states obtained without invoking this approximation. The CVS states are in general also not eigenstates of the full field-free Jacobian, and can lead to TD-EOM-CC populations that are non-stationary, complicating the interpretation of the TD-EOM-CC state. To ensure that the populations are stationary, we diagonalize the Jacobian $\mathbf{A}$ in the basis of all the CVS and non-CVS (valence) states by first constructing the Jacobian and overlap matrices

$$A_{ij} = \boldsymbol{L}_i \mathbf{A} \boldsymbol{R}_j, \quad S_{ij} = \boldsymbol{L}_i \boldsymbol{R}_j. \tag{17}$$

respectively in the reduced space. Assuming linear independence of the vectors in the basis, the solution of the generalized eigenvalue problem defined by $\mathbf{A}$ and $\mathbf{S}$ gives a new set of right and left eigenvectors of $\mathbf{A}$, which preserve populations when there is no interaction with the external field.

III. COMPUTATIONAL DETAILS

The electric field in Eq. (2) is represented as

$$\boldsymbol{\mathcal{E}}(t) = \boldsymbol{\mathcal{E}}_0 \cos(\omega_0(t - t_0) + \phi) f(t), \tag{18}$$

where $\boldsymbol{\mathcal{E}}_0$ is the peak electric field of the pulse in its polarization direction, ω_0 the carrier frequency and t_0 the central time of the pulse, and ϕ is the carrier-envelope phase. The envelope function $f(t)$ is chosen to have the Gaussian shape

$$f(t) = \begin{cases} e^{-(t-t_0)^2/(2\sigma^2)}, & -c \leq t \leq c, \\ 0, & \text{otherwise,} \end{cases} \tag{19}$$

where the RMS width is selected as $\sigma = 0.5$ and the envelope truncated at $c = 8\sigma$. In all calculations, we use the carrier-envelope phase $\phi = 0$ and the peak electric field strength $|\boldsymbol{\mathcal{E}}_0|$ of 10 a.u., which corresponds to the maximum intensity of $7.019 \times 10^{18}\,\mathrm{W\,cm^{-2}}$, calculated from the intensity relation $S_0 = |\boldsymbol{\mathcal{E}}_0|^2/Z_0$ where Z_0 is the impedance of free space.

All simulations are performed using a development version of the eT program [47] containing the TD-EOM-CC implementation described in Ref. [45]. The Runge-Kutta method known as RK4 is used to integrate Eq. (12) and Eq. (13), with time steps of 0.001 a.u. for neon, carbon monoxide, and p-aminophenol and 0.0001 a.u. for pyrrole.

IV. RESULTS AND DISCUSSION

A. Neon

In the following, the convergence properties of the final Raman-induced populations are investigated for the neon atom. This system is used for benchmarking purposes, as its small size allows for the use of larger basis sets. We focus on the convergence of the final population of the $B_v\,^1D$ valence-excited state, the lowest valence-excited state with a significant final population.

We first study the basis set convergence with respect to the cardinal number X of Dunning basis sets for CCS and CCSD levels of theory. The employed basis sets are cc-pVDZ, aug-cc-pVXZ (with X=D,...,6) and aug-cc-pCVXZ (with X=D,...,5). As the carrier frequency ω_0 of the electric field, we choose the average of two frequencies. The first frequency corresponds to the transition between the ground state X^1S and the $B_c\,^1P$ core-excited state. The second frequency corresponds to the transition between the $B_c\,^1P$ core-excited state and the $B_v\,^1D$ valence-excited state. The $B_c\,^1P$ and $B_v\,^1D$ states are chosen as they are, respectively, the lowest core-excited and valence excited states that get significantly populated in the Raman process, except for the cc-pVDZ basis set, where the order of $A_c\,^1S$ and $B_c\,^1P$ energy levels is inverted. In these calculations, we include 4 core-excited states and 12 valence-excited states. The frequencies used for the different basis sets and levels of theory are given in the Supporting Information.

From Fig. 2, we can observe how the final populations calculated with CCS and CCSD are considerably different, implying that CCS is not accurate enough to provide an adequate description of the system. The addition of functions for describing core correlation (aug-cc-pCVXZ) leads to slightly lower final populations compared to the corresponding basis sets without these functions (aug-cc-pVXZ). For CCSD, the results for 5Z and 6Z are very similar, implying that basis-set convergence is reached for 5Z. Continuing, the convergence of the final population of the $B_v\,^1D$ state is explored with respect to the number of valence- and core-excited states included in the calculation. The total of the probabilites of all de-
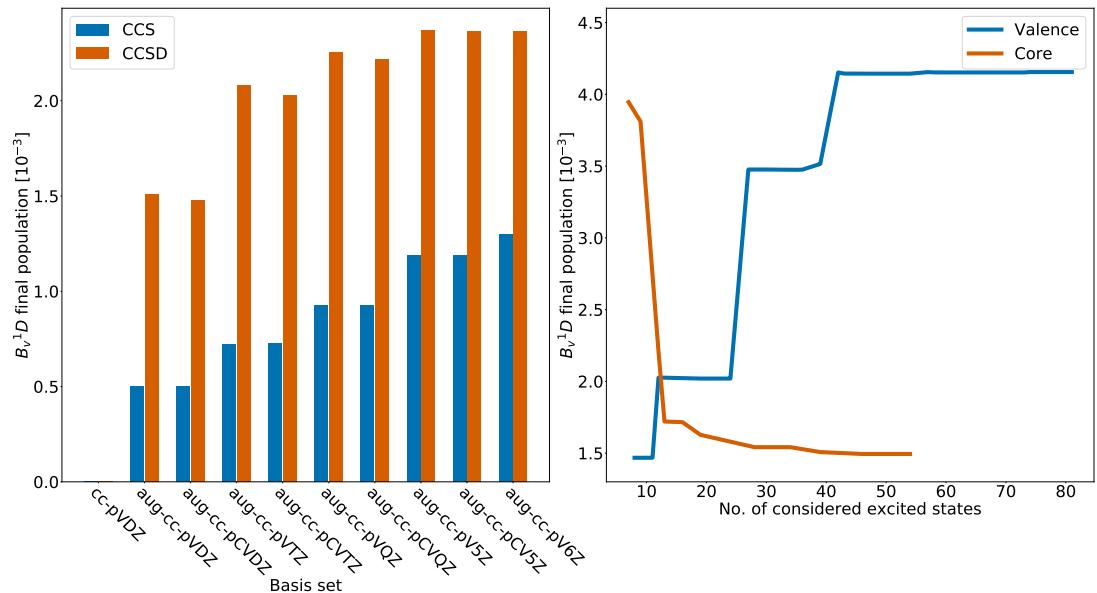
FIG. 2. The left panel shows the final population of the $B_v\,^1D$ states of neon for different choices of level of theory and basis set. The blue line in the right panel shows the final population of the same $B_v\,^1D$ states for different numbers of valence-excited states included in the simulation and the number of core-excited states fixed at 4, calculated with CCSD/aug-cc-pCVTZ. The red line in the right panel shows the final population of the same $B_v\,^1D$ states for different numbers of core-excited states included in the simulation and the number of valence-excited states fixed at 79, calculated with CCSD/aug-cc-pCVTZ.

generacies of a state is calculated, such that for instance the probabilities for the five degenerate states of D type are added together. We perform the calculations using the CCSD truncation level and the aug-cc-pCVTZ basis set. The right panel of Fig. 2 exhibits the convergence of the final population of the $B_v\,^1D$ states with respect to the number of valence-excited states included in the simulation, with the number of core-excited states fixed at 4. The results indicate that more than 40 valence-excited states are needed for convergence. An analogous procedure is performed, this time keeping the number of valence-excited states fixed while varying the number of core-excited states. In the right panel of Fig. 2, we can see how the final population of $B_v\,^1D$ starts to converge after around 15 core-excited states are included in the calculation.

B. CO

We continue by simulating ISXRS for the carbon monoxide molecule, which is linear and belongs to the $C_{\infty v}$ symmetry point group. Since the system is not centrally symmetric, results can differ depending on the polarization of the electric field. Theoretical and experimental studies of the core-excitation spectroscopy and ISXRS of this molecule have previously been car-

ried out. [48] In our simulations, the distance between the two nuclei is fixed at $1.128\,\text{Å}$, corresponding to the equilibrium bond length in the NIST database. [49] The internuclear axis of the molecule is aligned along the z-axis and the carbon atom is placed at the origin of the coordinate system while the oxygen atom is placed at $1.128\,\text{Å}$ along the z-axis. The carrier frequency of the external electric field is again chosen as the average between two frequencies. The first is the transition frequency between the ground state and the first core-excited state, which is the lowest-energy core-excited state that gets significantly populated during the Raman process. The second is the frequency of transition between this core-excited state and the third valence-excited state, which is the lowest valence-excited state that gets significantly populated. For CCS/aug-cc-pCVTZ, the frequency is $20.029\,089\,E_\text{h}$, while for CCSD/aug-cc-pCVTZ it is $19.504\,022\,E_\text{h}$, corresponding to the O K-edge.

To investigate transitions at the C K-edge, we choose the lowest-energy molecular orbital localized on the carbon atom as the molecular orbital used in the CVS approximation. The carrier frequency of the electric field is again chosen as average of the transition frequencies between the ground state and the lowest core-excited state that is significantly populated, and that between that core-excited state and the lowest valence-excited state that is significantly populated, resulting in a carrier fre-
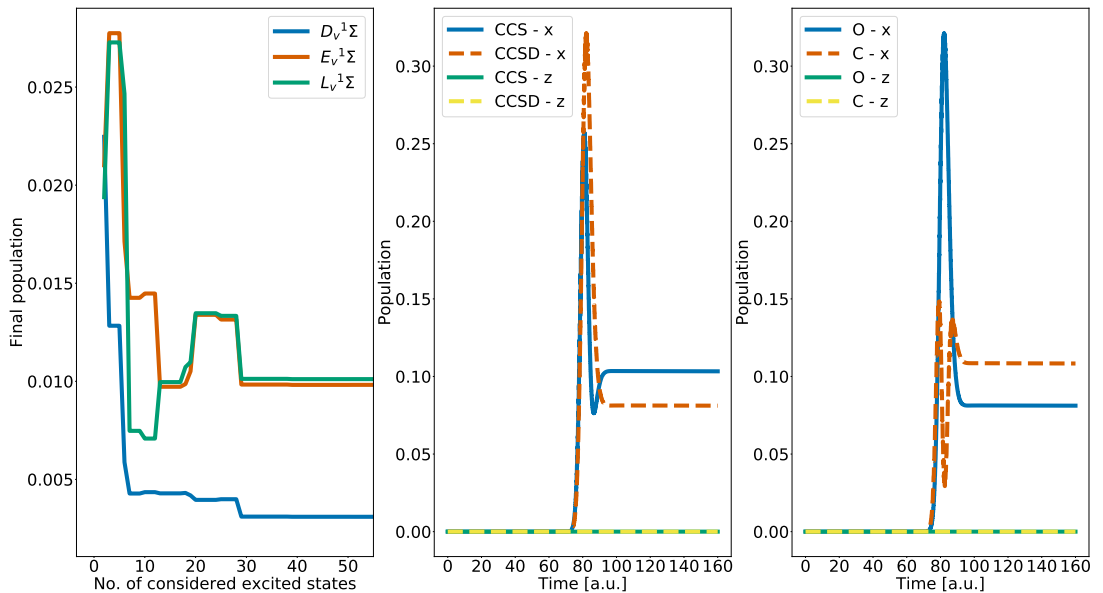
FIG. 3. The left panel shows the final population of the $D_v{}^1\Sigma$, $E_v{}^1\Sigma$, and $L_v{}^1\Sigma$ valence-excited state of carbon monoxide for different numbers of core-excited states included in the calculation, with the number of valence states fixed at 20 and the external electric field polarization in the positive z-direction. The central panel shows the time-dependent population of the third valence-excited state of carbon monoxide, calculated with the aug-cc-pCVTZ basis set and different levels of theory and electric field polarizations. The right panel shows the time-dependent population of the third valence-excited state for external electric fields tuned to different K-edges and with different polarizations, calculated with CCSD/aug-cc-pCVTZ.

quency of $10.402\,530\,E_\mathrm{h}$

In the carbon monoxide system, linearly polarized electric fields can be decomposed into two components: the polarization component parallel to the internuclear axis (along the z-axis) and the polarization component perpendicular to it (any direction in the xy-plane). As for neon, the convergence of the final population of certain valence-excited states is assessed with respect to the number of included core-excited states. The results are shown for the $D_v{}^1\Sigma$, $E_v{}^1\Sigma$ and $L_v{}^1\Sigma$ valence-excited states in the left panel of Fig. 3, demonstrating that convergence is attained by increasing the number of considered core-excited states. About 30 core-excited states are needed for convergence when the number of valence-excited states is fixed at 20. In the central panel of the figure, we can see how the time-dependent population of the third valence-excited state depends on the polarization of the electric field and level of theory, and also how the population is constant after the interaction with the field. The final population is exactly zero when the polarization is along the z-axis, as expected from the symmetry of the molecule and field. In the right panel, we can see how the time-dependent population of the third valence-excited state differs when the carrier frequency of the electric field is tuned to the K-edge of different elements (C or O). For the different tunings, the third

valence-excited state is reached through different transition pathways, involving other transition frequencies and transition moments. As for the results in the central panel, the population is exactly zero when the electric field is polarized along the z-axis, irrespective of the chosen frequency, for symmetry reasons. The populations are also constant after the interaction with the field

C. Pyrrole

We further increase the complexity of the modeled system by considering pyrrole, which belongs to the C_{2v} symmetry point group. The geometry of the molecule is obtained from the NIST database. [49], for which the molecule lies in the yz-plane and the symmetry axis is along the z-axis. The Supporting Information provides the geometry of the system, along with a figure that shows its orientation relative to the Cartesian coordinate axes. The final populations after the Raman process are assessed for the electric field polarization vector set equal to $(1, 0, 0)$ $(1, 1, 0)$ and $(1, 1, 1)$ in the chosen coordinate system. The Raman process involving the N K-edge is studied by performing calculations at the CCSD level of theory with aug-cc-pCVDZ for the nitrogen atom and aug-cc-pVDZ for the other atoms. The carrier frequency
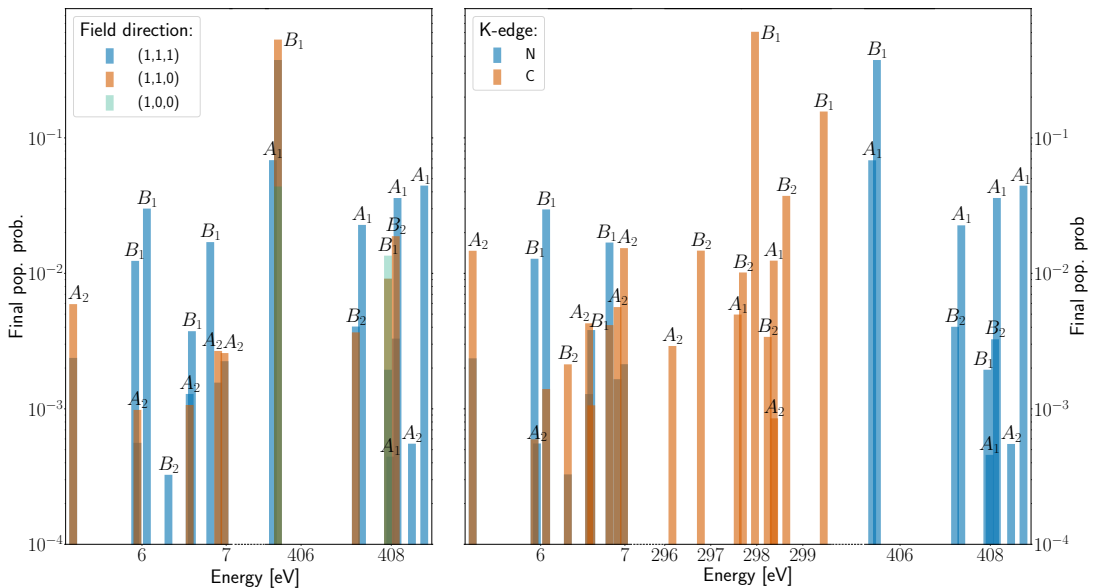
FIG. 4. The left panel displays the final populations of various excited states of pyrrole following ISXRS with different electric field polarizations, computed at the CCSD level of theory and with the aug-cc-pCVDZ basis set for the nitrogen atom and the aug-cc-pVDZ basis set for the other atoms. The right panel displays the final populations of different excited states of pyrrole for electric fields tuned to different K-edges, computed at the CCSD level of theory and with the aug-cc-pCVDZ basis set for the atom with the targeted K-edge shown in the inset and aug-cc-pVDZ basis set for the remaining atoms.

of the external electric field is chosen as the frequency of transition from the ground state to the most populated core-excited state, which is $14.901363\,E_{\mathrm{h}}$. The Raman process involving the C K-edge is studied by performing calculations at the CCSD level of theory with aug-cc-pCVDZ for the carbon atoms and aug-cc-pVDZ for the other atoms. The core-excited states are calculated by using the CVS approximation restricted to the molecular orbital with the second-lowest energy. The carrier frequency of the external electric field is set to $10.949885\,E_{\mathrm{h}}$, which is the transition frequency from the ground state to the fifth core-excited state, the lowest-energy core-excited state that is the most populated.

In the left panel of Fig. 4, we can see that new valence-excited states are populated as the polarization of the external electric field changes from $(1,0,0)$, to $(1,1,0)$, and to $(1,1,1)$. In particular, when the electric field is only polarized along the x-axis, there are no excitation to the valence-excited states. When the electric field has components along all three axes, all considered valence-excited states have a nonzero final population. An intermediate situation occurs when the electric field has components along both the x- and y-axes but not along the z-axis. This is since the different polarizations of the external electric field has components in different numbers of irreducible representation, enabling transitions to electronic states belonging to different irreducible repre-

sentations. In the right panel of Fig. 4, we can see how the final population of valence-excited states differs when the carrier frequency of the electric field is tuned to the N K-edge and C K-edge, calculated using the CVS approximation with the lowest- and next-to-lowest-energy molecular orbitals, respectively. In both cases, the polarization vector of the field is set to $(1,1,1)$. The valence-excited states that become populated are the same for the two K-edge frequencies, while the populations of the states are different.

D.   p-aminophenol

Finally, we consider the planar p-aminophenol molecule. The molecule belongs to the C_s symmetry point group, which only contains the mirror plane and the identity as symmetry elements. This molecule is chosen in order to investigate if charge migration between the functional groups located at the opposite side of the aromatic ring can be observed, as the electronic charge can easily travel along the aromatic electron cloud. [50]

Compared to the systems analyzed previously, which offer only limited potential for charge migration due to their small sizes, the p-aminophenol molecule is a larger system containing two strongly electron donor substituents (amino and hydroxyl) on a benzene ring. [51].
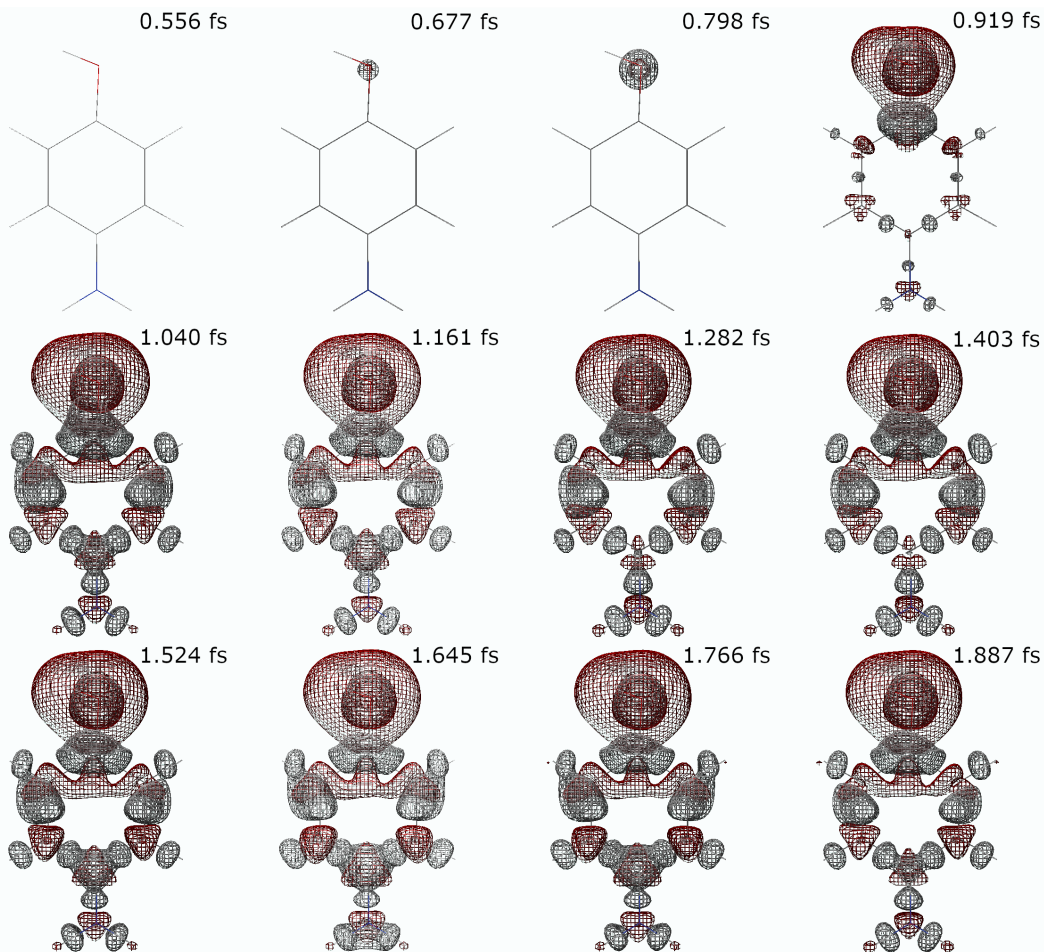
FIG. 5. Positive (gray) and negative (red) electronic isodensity surfaces of the time-dependent density after subtracting the ground state density of p-aminophenol, at the times specified at the top right corner of each subfigure. The structure of the p-aminophenol molecule is also shown in each subfigure.

We can thus expect a localized excitation to be followed by long-range charge migration.

The geometry of p-aminophenol is calculated at the B3LYP/aug-cc-pVDZ level of theory, and the molecule is placed in the xy-plane. The Supporting Information includes the geometry and a figure that illustrates the orientation of the molecule relative to the Cartesian coordinate axes. For the subsequent calculations, aug-cc-pCVDZ is used for the oxygen atom and aug-cc-pVDZ for all other atoms. The carrier frequency is chosen as $19.883\,479\,E_{\mathrm{h}}$, which corresponds to the frequency of transition from the ground state to the fourth core-excited state, which is the most populated state among the two lowest-energy core-excited states that have a non-

zero population after the Raman process.

In Fig. 5, the charge migration is illustrated through isodensity surfaces of the time-dependent density after subtracting the ground state density, calculated at different points in time. After the interaction with the external electromagnetic pulse, we can observe how the core excitation of the oxygen atom is reflected in a positive charge arising around that nucleus, enclosed in a negatively charged region at a bigger distance from the oxygen nucleus. This is followed by an alternating pattern of regions with increased or decreased electronic charge throughout the entire benzene ring up to the nitrogen atom of the amino group. In particular, the atoms of the ring gain some negative charge while the bonds be-

come more positively charged, and the bonds are thus expected to be weakened. Finally, we can observe how the nitrogen atom becomes negatively charged. As predicted, we observe a localized excitation at the hydroxyl substituent following oxygen K-edge excitation, followed by long-range charge migration, in accordance with what one could expect from a superposition of valence-excited states generated by ISXRS.

In the supplemental material we have included a movie that shows the temporal evolution of the electronic density depicted through isodensity surfaces of the time-dependent density difference, illustrating how the density oscillates after the interaction with the external electric field. The generation of electronic wavepackets with external laser pulses is interesting from an experimental point of view, as it represents the first step of controlling chemical reactions with laser pulses.

V.   CONCLUSION

In this work, a time-dependent equation-of-motion coupled cluster model of ISXRS has been presented. First, we assessed the convergence of the final population of neon valence states with respect to different calculation parameters: the level of coupled cluster theory, the choice of basis set, and choices of the total number of valence- and core-excited states. We observed how the adequate description of the system required a proper representation of correlation and a sufficiently flexible basis set, since the CCS level of theory and basis sets without augmentation performed poorly. We also demonstrated that convergence of the population of a valence-excited state of neon was achieved when increasing the number of valence- and core-excited states for the given level of theory and basis set. Subsequently, the final populations of carbon monoxide states were assessed with respect to the number of included core-excited states. The results showed convergence for several valence-excited states for the given level of theory and basis set.

Furthermore, we demonstrated that the final populations of states of both carbon monoxide and pyrrole are significantly affected by the polarization of the external electric field, as symmetry can enable and forbid the transition to some of the excited states within the bandwidth of the pulse. We also assessed how the results were affected by tuning the external electric field to the K-edge of the different atoms, where the frequencies were calculated with the CVS approximation targeting the core molecular orbitals of the atoms. We observed how a different choice of K-edge led to changes in final populations as the final states were reached through different transition pathways.

After investigating ISXRS by neon, carbon monoxide, and pyrrole, we studied the time evolution of the electronic density of p-aminophenol. The ground-state density was subtracted from the time-dependent density, and the density difference was visualized through isodensity surfaces in real space. We observed the rapid formation of a valence wavepacket and subsequent charge migration in the molecule. Simulations of field-induced charge migration in molecular systems can be used to predict how chemical reactions can be controlled by external electric fields, which we believe will be a subject of further interest in the near future.

ACKNOWLEDGMENTS

[1] C. Pellegrini, A. Marinelli, and S. Reiche, The physics of x-ray free-electron lasers, Rev. Mod. Phys. **88**, 015006 (2016).

[2] C. Pellegrini, The development of XFELs, Nat. Rev. Phys. **2**, 330 (2020).

[3] A. McPherson, G. Gibson, H. Jara, U. Johann, T. S. Luk, I. McIntyre, K. Boyer, and C. K. Rhodes, Studies of multiphoton production of vacuum-ultraviolet radiation in the rare gases, J. Opt. Soc. Am. B **4**, 595 (1987).

[4] M. Ferray, A. L'Huillier, X. Li, L. Lompre, G. Mainfray, and C. Manus, Multiple-harmonic conversion of 1064 nm radiation in rare gases, J. Phys. B: At. Mol. Opt. Phys. **21**, L31 (1988).

[5] T. Gaumnitz, A. Jain, Y. Pertot, M. Huppert, I. Jordan, F. Ardana-Lamas, and H. J. Wörner, Streaking of 43-attosecond soft-x-ray pulses generated by a passively cep-stable mid-infrared driver, Opt. Express **25**, 27506 (2017).

[6] S. R. Leone, C. W. McCurdy, J. Burgdörfer, L. S. Cederbaum, Z. Chang, N. Dudovich, J. Feist, C. H. Greene, M. Ivanov, R. Kienberger, U. Keller, M. F. Kling, Z.-H. Loh, T. Pfeifer, A. N. Pfeiffer, R. Santra, K. Schafer, A. Stolow, U. Thumm, and M. J. J. Vrakking, What will it take to observe processes in 'real time'?, Nat. Photonics **8**, 162 (2014).

[7] A. Baltuška, T. Udem, M. Ulberacker, M. Hentschel, E. Goulielmakis, C. Gohle, R. Holzwarth, V. S. Yakovlev, A. Scrinzi, T. W. Hänsch, and F. Krausz, Attosecond control of electronic processes by intense light fields, Nature **421**, 611 (2003).

[8] J. Duris, S. Li, T. Driver, E. G. Champenois, J. P. MacArthur, A. A. Lutman, Z. Zhang, P. Rosenberger, J. W. Aldrich, R. Coffee, G. Coslovich, F.-J. Decker, J. M. Glownia, G. Hartmann, W. Helml, A. Kamalov,

J. Knurr, J. Krzywinski, M.-F. Lin, J. P. Marangos, M. Nantel, A. Natan, J. T. O'Neal, N. Shivaram, P. Walter, A. L. Wang, J. J. Welch, J. Z. Wolf, Thomas J. A. Xu, M. F. Kling, P. H. Bucksbaum, A. Zholents, Z. Huang, J. P. Cryan, and A. Marinelli, Tunable isolated attosecond x-ray pulses with gigawatt peak power from a free-electron laser, Nat. Photonics **14**, 30 (2020).

[9] M. Drescher, M. Hentschel, R. Kienberger, G. Tempea, C. Spielmann, G. A. Reider, P. B. Corkum, and F. Krausz, X-ray pulses approaching the attosecond frontier, Science **291**, 1923 (2001).

[10] S. Li, T. Driver, P. Rosenberger, E. G. Champenois, J. Duris, A. Al-Haddad, V. Averbukh, J. C. T. Barnard, N. Berrah, C. Bostedt, P. H. Bucksbaum, R. N. Coffee, L. F. DiMauro, L. Fang, D. Garratt, A. Gatton, Z. Guo, G. Hartmann, D. Haxton, W. Helml, Z. Huang, A. C. LaForge, A. Kamalov, J. Knurr, M.-F. Lin, A. A. Lutman, J. P. MacArthur, J. P. Marangos, M. Nantel, A. Natan, R. Obaid, J. T. O'Neal, N. H. Shivaram, A. Schori, P. Walter, A. L. Wang, T. J. A. Wolf, Z. Zhang, M. F. Kling, A. Marinelli, and J. P. Cryan, Attosecond coherent electron motion in auger-meitner decay, Science **375**, 285 (2022).

[11] F. Calegari, D. Ayuso, A. Trabattoni, L. Belshaw, S. D. Camillis, S. Anumula, F. Frassetto, L. Poletto, A. Palacios, P. Decleva, J. B. Greenwood, F. Martín, and M. Nisoli, Ultrafast electron dynamics in phenylalanine initiated by attosecond pulses, Science **346**, 336 (2014).

[12] P. M. Kraus, B. Mignolet, D. Baykusheva, A. Rupenyan, L. Horný, E. F. Penka, G. Grassi, O. I. Tolstikhin, J. Schneider, F. Jensen, L. B. Madsen, A. D. Bandrauk, F. Remacle, and H. J. Wörner, Measurement and laser control of attosecond charge migration in ionized iodoacetylene, Science **350**, 790 (2015).

[13] J. D. Biggs, Y. Zhang, D. Healion, and S. Mukamel, Watching energy transfer in metalloporphyrin heterodimers using stimulated x-ray raman spectroscopy, Proc. Natl. Acad. Sci. **110**, 15597 (2013).

[14] S. Tanaka and S. Mukamel, Probing exciton dynamics using raman resonances in femtosecond x-ray four-wave mixing, Phys. Rev. A **67**, 033818 (2003).

[15] N. Rohringer, X-ray raman scattering: a building block for nonlinear spectroscopy, Philos. Trans. Royal Soc. A: Math. Phys. Eng. Sci. **377**, 20170471 (2019).

[16] J. D. Biggs, Y. Zhang, D. Healion, and S. Mukamel, Two-dimensional stimulated resonance raman spectroscopy of molecules with broadband x-ray pulses, J. Chem. Phys. **136**, 174117 (2012).

[17] C. Weninger, M. Purvis, D. Ryan, R. A. London, J. D. Bozek, C. Bostedt, A. Graf, G. Brown, J. J. Rocca, and N. Rohringer, Stimulated electronic x-ray raman scattering, Phys. Rev. Lett. **111**, 233902 (2013).

[18] J. T. O'Neal, E. G. Champenois, S. Oberli, R. Obaid, A. Al-Haddad, J. Barnard, N. Berrah, R. Coffee, J. Duris, G. Galinis, D. Garratt, J. M. Glownia, D. Haxton, P. Ho, S. Li, X. Li, J. MacArthur, J. P. Marangos, A. Natan, N. Shivaram, D. S. Slaughter, P. Walter, S. Wandel, L. Young, C. Bostedt, P. H. Bucksbaum, A. Picón, A. Marinelli, and J. P. Cryan, Electronic population transfer via impulsive stimulated x-ray raman scattering with attosecond soft-x-ray pulses, Phys. Rev. Lett. **125**, 073203 (2020).

[19] X. Li, N. Govind, C. Isborn, A. E. DePrince III, and K. Lopata, Real-time time-dependent electronic struc-

ture theory, Chem. Rev. **120**, 9951 (2020).

[20] S. Kvaal, Ab initio quantum dynamics using coupled-cluster, J. Chem. Phys. **136**, 194109 (2012).

[21] C. Huber and T. Klamroth, Explicitly time-dependent coupled cluster singles doubles calculations of laser-driven many-electron dynamics, J. Chem. Phys. **134**, 054113 (2011).

[22] D. A. Pigg, G. Hagen, H. Nam, and T. Papenbrock, Time-dependent coupled-cluster method for atomic nuclei, Phys. Rev. C **86**, 014308 (2012).

[23] A. S. Skeidsvoll, A. Balbi, and H. Koch, Time-dependent coupled-cluster theory for ultrafast transient-absorption spectroscopy, Phys. Rev. A **102**, 023115 (2020).

[24] Y. C. Park, A. Perera, and R. J. Bartlett, Equation of motion coupled-cluster for core excitation spectra: Two complementary approaches, J. Chem. Phys. **151**, 164117 (2019).

[25] H. E. Kristiansen, Ø. S. Schøyen, S. Kvaal, and T. B. Pedersen, Numerical stability of time-dependent coupled-cluster methods for many-electron dynamics in intense laser pulses, J. Chem. Phys. **152**, 071102 (2020).

[26] T. B. Pedersen, H. E. Kristiansen, T. Bodenstein, S. Kvaal, and Ø. S. Schøyen, Interpretation of coupled-cluster many-electron dynamics in terms of stationary states, J. Chem. Theory Comput. **17**, 388 (2021).

[27] Y. C. Park, A. Perera, and R. J. Bartlett, Equation of motion coupled-cluster study of core excitation spectra ii: Beyond the dipole approximation, J. Chem. Phys. **155**, 094103 (2021).

[28] T. B. Pedersen and S. Kvaal, Symplectic integration and physical interpretation of time-dependent coupled-cluster theory, J. Chem. Phys. **150**, 144106 (2019).

[29] H. Koch and P. Jørgensen, Coupled cluster response functions, J. Chem. Phys. **93**, 3333 (1990).

[30] T. B. Pedersen and H. Koch, Coupled cluster response functions revisited, J. Chem. Phys. **106**, 8059 (1997).

[31] D. R. Nascimento and A. E. DePrince III, Linear absorption spectra from explicitly time-dependent equation-of-motion coupled-cluster theory, J. Chem. Theory Comput. **12**, 5834 (2016).

[32] H. B. Schlegel, S. M. Smith, and X. Li, Electronic optical response of molecules in intense fields: Comparison of td-hf, td-cis, and td-cis (d) approaches, J. Chem. Phys. **126**, 244110 (2007).

[33] E. Luppi and M. Head-Gordon, Computation of high-harmonic generation spectra of h2 and n2 in intense laser pulses using quantum chemistry methods and time-dependent density functional theory, Mol. Phys. **110**, 909 (2012).

[34] D. R. Nascimento and A. E. DePrince, Simulation of near-edge x-ray absorption fine structure with time-dependent equation-of-motion coupled-cluster theory, J. Phys. Chem. Lett. **8**, 2951 (2017).

[35] D. R. Nascimento and A. E. DePrince III, A general time-domain formulation of equation-of-motion coupled-cluster theory for linear spectroscopy, J. Chem. Phys. **151**, 204107 (2019).

[36] L. N. Koulias, D. B. Williams-Young, D. R. Nascimento, A. E. DePrince, and X. Li, Relativistic real-time time-dependent equation-of-motion coupled-cluster, J. Chem. Theory Comput. **15**, 6617 (2019).

[37] A. S. Skeidsvoll and H. Koch, Comparing real-time coupled cluster methods through simulation of collective rabi oscillations (2023), arXiv:2301.05546 [physics.chem-ph].

[38] D. Mukherjee and P. Mukherjee, A response-function approach to the direct calculation of the transition-energy in a multiple-cluster expansion formalism, Chem. Phys. **39**, 325 (1979).

[39] K. Emrich, An extension of the coupled cluster formalism to excited states (i), Nucl. Phys. A **351**, 379 (1981).

[40] J. F. Stanton and R. J. Bartlett, The equation of motion coupled-cluster method. a systematic biorthogonal approach to molecular excitation energies, transition probabilities, and excited state properties, J. Chem. Phys. **98**, 7029 (1993).

[41] P. Norman and A. Dreuw, Simulating x-ray spectroscopies and calculating core-excited states of molecules, Chem. Rev. **118**, 7208 (2018).

[42] L. S. Cederbaum, W. Domcke, and J. Schirmer, Many-body theory of core holes, Phys. Rev. A **22**, 206 (1980).

[43] S. Coriani and H. Koch, Communication: X-ray absorption spectra and core-ionization potentials within a core-valence separated coupled cluster framework, J. Chem. Phys. **143**, 181103 (2015).

[44] T. Helgaker, P. Jørgensen, and J. Olsen, Coupled-cluster theory, in *Molecular Electronic-Structure Theory* (John Wiley & Sons, Ltd, 2000) Chap. 13, pp. 648–723.

[45] A. S. Skeidsvoll, T. Moitra, A. Balbi, A. C. Paul, S. Coriani, and H. Koch, Simulating weak-field attosecond processes with a lanczos reduced basis approach to time-dependent equation-of-motion coupled-cluster theory, Phys. Rev. A **105**, 023103 (2022).

[46] A. Barth and L. S. Cederbaum, Many-body theory of core-valence excitations, Phys. Rev. A **23**, 1038 (1981).

[47] S. D. Folkestad, E. F. Kjønstad, R. H. Myhre, J. H. Andersen, A. Balbi, S. Coriani, T. Giovannini, L. Goletto, T. S. Haugland, A. Hutcheson, I.-M. Høyvik, T. Moitra, A. C. Paul, M. Scavino, A. S. Skeidsvoll, Å. H. Tveten, and H. Koch, et 1.0: An open source electronic structure program with emphasis on coupled cluster and multilevel methods, J. Chem. Phys. **152**, 184103 (2020).

[48] R. Püttner, I. Dominguez, T. J. Morgan, C. Cisneros, R. F. Fink, E. Rotenberg, T. Warwick, M. Domke, G. Kaindl, and A. S. Schlachter, Vibrationally resolved o 1s core-excitation spectra of co and no, Phys. Rev. A **59**, 3415 (1999).

[49] F. Lovas, E. Tiemann, J. Coursey, S. Kotochigova, J. Chang, K. Olsen, and R. Dragoset, Diatomic spectral database (2003).

[50] H. J. Wörner, C. A. Arrell, N. Banerji, A. Cannizzo, M. Chergui, A. K. Das, P. Hamm, U. Keller, P. M. Kraus, E. Liberatore, P. Lopez-Tarifa, M. Lucchini, M. Meuwly, C. Milne, J.-E. Moser, U. Rothlisberger, G. Smolentsev, J. Teuscher, J. A. van Bokhoven, and O. Wenger, Charge migration and charge transfer in molecular systems, Struct. Dyn. **4**, 061508 (2017).

[51] T. Brinck, M. Haeberlein, and M. Jonsson, A computational analysis of substituent effects on the o- h bond dissociation energy in phenols: polar versus radical effects, J. Am. Chem. Soc. **119**, 4239 (1997).

NTNU

Norwegian University of
Science and Technology