



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

Skeidsvoll, Andreas S.; Moitra, Torsha; Balbi, Alice; Paul, Alexander C.; Coriani, Sonia; Koch, Henrik

*Published in:*  
Physical Review A

*Link to article, DOI:*  
[10.1103/PhysRevA.105.023103](https://doi.org/10.1103/PhysRevA.105.023103)

*Publication date:*  
2022

*Document Version*  
Peer reviewed version

[Link back to DTU Orbit](#)

*Citation (APA):*  
Skeidsvoll, A. S., Moitra, T., Balbi, A., Paul, A. C., Coriani, S., & Koch, H. (2022). Simulating weak-field attosecond processes with a Lanczos reduced basis approach to time-dependent equation-of-motion coupled-cluster theory. *Physical Review A*, 105(2), Article 023103. <https://doi.org/10.1103/PhysRevA.105.023103>

---

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

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

Andreas S. Skeidsvoll,<sup>1,\*</sup> Torsha Moitra,<sup>2,\*</sup> Alice Balbi,<sup>3</sup>  
Alexander C. Paul,<sup>1</sup> Sonia Coriani,<sup>2,1,†</sup> and Henrik Koch<sup>1,3,‡</sup>

<sup>1</sup>*Department of Chemistry, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway*

<sup>2</sup>*DTU Chemistry—Department of Chemistry, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark*

<sup>3</sup>*Scuola Normale Superiore, Piazza dei Cavalieri, 7, IT-56126 Pisa, PI, Italy*

(Dated: October 27, 2021)

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 ultra-short 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 (TDCCSD) 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 (CVS) 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 ( $1 \times 10^{-18}$  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 real-time ultrafast dynamics and phenomena in both perturbative and non-perturbative regimes. [7, 8]

The development of real-time methods commenced in the late 70s and early 80s in the field of nuclear physics [9–11]. Despite these early endeavours, 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].

In this work, we present an implementation and representative case studies of the TD-EOM-CC model for sim-

ulating 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 TDCC counterpart in weak fields, at significantly lower computational costs. The reduction in cost enables the study of larger systems.

The article is organized as follows. In Section 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 Section III. In Section 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 scheme is tested for LiH. Then, two-photon absorption phenomenon has been captured using a stepwise procedure emulating the actual physical process for  $C_2H_4$ . Finally, we have 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 Section V.

## II. THEORY

### A. System

We model the system, comprising a molecule and its interaction with laser pulses, with the Hamiltonian

$$H(t) = H^{(0)} + V(t) , \quad (1)$$

where the field-free Hamiltonian  $H^{(0)}$  describes the molecule with fixed nuclei and without interactions with the external electromagnetic field. The semi-classical

\* These authors contributed equally

† Electronic mail: [soco@kemi.dtu.dk](mailto:soco@kemi.dtu.dk)

‡ Electronic mail: [henrik.koch@sns.it](mailto:henrik.koch@sns.it)

time-dependent interaction term, written in the dipole approximation and length gauge, is

$$V(t) = -\mathbf{d} \cdot \mathcal{E}(t), \quad (2)$$

and describes the interaction between the molecular electrons and the external electromagnetic field. The latter is represented by the electric field  $\mathcal{E}(t) = [\mathcal{E}_x(t) \ \mathcal{E}_y(t) \ \mathcal{E}_z(t)]^T$  and electronic dipole operator vectors,  $\mathbf{d} = [d_x \ d_y \ d_z]^T$ . 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,

$$\mathcal{E}(t) = \sum_n \mathcal{E}_{0n} \cos(\omega_{0n}(t - t_{0n}) + \phi_n) f_n(t). \quad (3)$$

The field of laser pulse  $n$  has an associated carrier frequency  $\omega_{0n}$ , peak strength  $|\mathcal{E}_{0n}|$  and polarization  $\mathcal{E}_{0n}/|\mathcal{E}_{0n}|$ , and an  $8\sigma_n$ -truncated Gaussian envelope function

$$f_n(t) = \begin{cases} e^{-(t-t_{0n})^2/(2\sigma_n^2)}, & |t - t_{0n}| \leq 8\sigma_n, \\ 0 & \text{otherwise,} \end{cases} \quad (4)$$

with duration specified by  $\sigma_n$ , the temporal root-mean-square (RMS) width. It is also specified by the central time  $t_{0n}$  and the carrier-envelope phase (CEP)  $\phi_n$ . 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]

$$\Delta E = \int_0^\infty \omega S(\omega) d\omega, \quad (5)$$

where  $S(\omega)$  is the response function

$$S(\omega) = -2 \text{Im}[\langle \tilde{\mathbf{d}} \rangle(\omega) \cdot \tilde{\mathcal{E}}^*(\omega)], \quad \omega > 0. \quad (6)$$

The vectors  $\langle \tilde{\mathbf{d}} \rangle(\omega)$  and  $\tilde{\mathcal{E}}(\omega)$  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  $\omega$ , respectively [34].

## B. TD-EOM-CC states

The time-dependent ket and bra of a TD-EOM-CC state can be expressed as

$$|\Psi(t)\rangle = \sum_i |\psi_i\rangle s_i(t), \quad \langle \tilde{\Psi}(t)| = \sum_i k_i(t) \langle \tilde{\psi}_i|, \quad (7)$$

where the italic indices  $i, j$  are used to denote general EOM-CC states, including the ground state with index

0. The time-independent EOM-CC kets and bras are given by

$$|\psi_i\rangle = e^T R_i |\text{HF}\rangle, \quad \langle \tilde{\psi}_i| = \langle \text{HF}| L_i e^{-T}. \quad (8)$$

We assume that the EOM-CC states are biorthonormal

$$\langle \tilde{\psi}_i | \psi_j \rangle = \delta_{ij}. \quad (9)$$

In the following, we let the indices  $\kappa, \lambda$  denote general determinants in the projection space, including the reference Hartree-Fock determinant with index 0. We use the indices  $\mu, \nu$ , on the other hand, to denote excited determinants.

The cluster operator  $T$  and the right and left operators  $R_i, L_i$  of Eq. (8) can be expressed as linear expansions in a finite set of operators  $\tau_\kappa, \tau_\kappa^\dagger$ ,

$$T = \sum_\mu \tau_\mu t_\mu, \quad R_i = \sum_\kappa \tau_\kappa r_{\kappa i}, \quad L_i = \sum_\kappa l_{i\kappa} \tau_\kappa^\dagger, \quad (10)$$

where the operator with index 0 is the unit operator,

$$\tau_0 = \tau_0^\dagger = 1, \quad (11)$$

and the  $\tau_\mu$  and  $\tau_\mu^\dagger$  operators generate excited determinants from the ket and bra reference Hartree-Fock determinants, respectively,

$$\tau_\mu |\text{HF}\rangle = |\mu\rangle, \quad \langle \text{HF}| \tau_\mu^\dagger = \langle \mu|, \quad (12)$$

$$\tau_\mu^\dagger |\text{HF}\rangle = 0, \quad \langle \text{HF}| \tau_\mu = 0. \quad (13)$$

We assume that the determinants are biorthogonal,

$$\langle \kappa | \lambda \rangle = \delta_{\kappa\lambda}. \quad (14)$$

If all possible electronic excitations are included in the summations in Eq. (10), the method is equivalent to full configuration interaction (FCI). 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 (CCSD) method, where summation is only done over single and double excitations. We do not explicitly state the excitation levels included in following expressions, since they hold for both restricted and unrestricted summation.

The cluster amplitudes  $t_\mu$  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

$$\langle \mu | \bar{H}^{(0)} | \text{HF} \rangle = 0, \quad (15)$$

where the similarity transformation of an operator  $X$  is denoted by an overbar,

$$\bar{X} = e^{-T} X e^T. \quad (16)$$

After the optimal cluster amplitudes  $t_\mu$  have been determined, the right and left vectors of EOM-CC state  $i$ , with components  $r_{\kappa i}$  and  $l_{i\kappa}$ , can be found as right and left eigenvectors of the field-free Hamiltonian matrix, with elements

$$H_{\kappa\lambda}^{(0)} = \langle \kappa | \bar{H}^{(0)} | \lambda \rangle . \quad (17)$$

The right and left eigenvectors of the matrix in Eq. (17) with the lowest eigenvalue, specifying the ground EOM-CC state with index 0, have the following structure

$$r_{00} = 1 , \quad r_{\mu 0} = 0 , \quad (18)$$

$$l_{00} = 1 , \quad l_{0\mu} = \bar{t}_\mu . \quad (19)$$

The multipliers  $\bar{t}_\mu$  are solutions to the equations

$$\langle \text{HF} | \bar{H}^{(0)} | \nu \rangle + \sum_\mu \bar{t}_\mu A_{\mu\nu}^{(0)} = 0 , \quad (20)$$

where elements of the field-free coupled-cluster Jacobian matrix  $\mathbf{A}^{(0)}$  are given by

$$A_{\mu\nu}^{(0)} = \langle \mu | [\bar{H}^{(0)}, \tau_\nu] | \text{HF} \rangle . \quad (21)$$

The other right and left eigenvectors of the matrix in Eq. (17) correspond to excited EOM-CC states, denoted by the italic indices  $m, n$ . The eigenvectors have the following reference determinant components

$$r_{0m} = - \sum_\mu \bar{t}_\mu r_{\mu m} , \quad (22)$$

$$l_{m0} = 0 . \quad (23)$$

These components enforce the biorthogonality between the ground and excited states, in accordance with Eq. (9). The vectors  $\mathbf{R}_m$  and  $\mathbf{L}_m$ , containing the components  $r_{\mu m}$  and  $l_{m\mu}$  of excited EOM-CC state  $m$ , are right and left eigenvectors of  $\mathbf{A}^{(0)}$ .

### 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)

$$\iota \frac{\partial}{\partial t} |\Psi(t)\rangle = H(t) |\Psi(t)\rangle , \quad (24)$$

where  $\iota$  denotes the imaginary unit, onto the bra of EOM-CC state  $i$ , giving

$$\iota \frac{\partial s_i(t)}{\partial t} = \sum_j H_{ij}(t) s_j(t) , \quad (25)$$

where the matrix elements of an operator  $X(t)$  are given by

$$X_{ij}(t) = \langle \tilde{\psi}_i | X(t) | \psi_j \rangle . \quad (26)$$

Likewise, the time derivative of the coefficients of the TD-EOM-CC bra can be found from projecting the bra TDSE,

$$-\iota \frac{\partial}{\partial t} \langle \tilde{\Psi}(t) | = \langle \tilde{\Psi}(t) | H(t) \quad (27)$$

onto the ket of EOM-CC state  $j$ , giving

$$-\iota \frac{\partial k_j(t)}{\partial t} = \sum_i k_i(t) H_{ij}(t) . \quad (28)$$

To our knowledge, Eq. (25) and Eq. (28) were first presented in Ref. [18], and were also used in Ref. [19]. After time-dependent coefficients have been obtained from solution of these equations, the time-dependent expectation value of an operator  $X(t)$  can be calculated according to

$$\langle X \rangle(t) = \sum_{ij} k_i(t) X_{ij}(t) s_j(t) . \quad (29)$$

### D. Asymmetric band Lanczos

The asymmetric band Lanczos algorithm is a generalization of the simple asymmetric Lanczos algorithm, employing multiple starting vectors instead of single ones [35–37]. The asymmetric square matrix  $\mathbf{M}$  and the sequences  $(\mathbf{b}_1, \dots, \mathbf{b}_m)$  and  $(\mathbf{c}_1, \dots, \mathbf{c}_p)$  of right and left starting vectors, respectively, determine the following sequences of vectors

$$\begin{aligned} & (\beta_i(\mathbf{M}, \mathbf{b}_1, \dots, \mathbf{b}_m))_{i=1}^j \\ & = \underbrace{(\mathbf{b}_1, \dots, \mathbf{b}_m, \mathbf{M}\mathbf{b}_1, \dots, \mathbf{M}\mathbf{b}_m, \mathbf{M}^2\mathbf{b}_1, \dots)}_j , \end{aligned} \quad (30)$$

$$\begin{aligned} & (\gamma_i(\mathbf{M}^T, \mathbf{c}_1, \dots, \mathbf{c}_p))_{i=1}^j \\ & = \underbrace{(\mathbf{c}_1, \dots, \mathbf{c}_p, \mathbf{M}^T\mathbf{c}_1, \dots, \mathbf{M}^T\mathbf{c}_p, (\mathbf{M}^T)^2\mathbf{c}_1, \dots)}_j , \end{aligned} \quad (31)$$

which can be linearly dependent. We define the right and left band Krylov subspaces as

$$\begin{aligned} & \mathcal{K}_j(\mathbf{M}, \mathbf{b}_1, \dots, \mathbf{b}_m) \\ & = \text{span} \{ (\beta_i(\mathbf{M}, \mathbf{b}_1, \dots, \mathbf{b}_m))_{i=1}^j \} , \end{aligned} \quad (32)$$

$$\begin{aligned} & \mathcal{K}_j(\mathbf{M}^T, \mathbf{c}_1, \dots, \mathbf{c}_p) \\ & = \text{span} \{ (\gamma_i(\mathbf{M}^T, \mathbf{c}_1, \dots, \mathbf{c}_p))_{i=1}^j \} , \end{aligned} \quad (33)$$

of dimensions  $n \leq j$ . Note that these subspaces can also be regarded as block Krylov subspaces [37] whenever  $j$  is a multiple of the number of start vectors.

The purpose of the asymmetric band Lanczos algorithm is to iteratively generate  $n$  linearly independent right and left Lanczos vectors,  $(\mathbf{v}_1, \dots, \mathbf{v}_n)$  and  $(\mathbf{w}_1, \dots, \mathbf{w}_n)$ , forming bases for the  $n$ -dimensional Krylov subspaces  $\mathcal{K}_j(\mathbf{M}, \mathbf{b}_1, \dots, \mathbf{b}_m)$  and  $\mathcal{K}_j(\mathbf{M}^T, \mathbf{c}_1, \dots, \mathbf{c}_p)$ ,

respectively. The  $n$  right and left Lanczos vectors can be generated with the recurrence relations [35]

$$\mathbf{M}\mathbf{V}_n = \mathbf{V}_n\mathbf{T}_n + \hat{\mathbf{V}}_n^{\text{C}} + \hat{\mathbf{V}}_n^{\text{D}}, \quad (34)$$

$$\mathbf{M}^T\mathbf{W}_n = \mathbf{W}_n\tilde{\mathbf{T}}_n + \hat{\mathbf{W}}_n^{\text{C}} + \hat{\mathbf{W}}_n^{\text{D}}, \quad (35)$$

where the right and left Lanczos vectors form the matrices  $\mathbf{V}_n = [\mathbf{v}_1 \cdots \mathbf{v}_n]$  and  $\mathbf{W}_n = [\mathbf{w}_1 \cdots \mathbf{w}_n]$ , respectively. The unnormalized vectors that form the nonzero columns of  $\hat{\mathbf{V}}_n^{\text{C}} = [\mathbf{0} \cdots \mathbf{0} \hat{\mathbf{v}}_{n+1} \cdots \hat{\mathbf{v}}_{n+m_c}]$  and  $\hat{\mathbf{W}}_n^{\text{C}} = [\mathbf{0} \cdots \mathbf{0} \hat{\mathbf{w}}_{n+1} \cdots \hat{\mathbf{w}}_{n+p_c}]$  serve as candidates for the next right and left Lanczos vectors, respectively. The sparse matrices  $\hat{\mathbf{V}}_n^{\text{D}}$  and  $\hat{\mathbf{W}}_n^{\text{D}}$  contain unnormalized candidates from previous iterations that have been deflated (i.e. rejected) due to linear dependence on already accepted right and left Lanczos vectors, respectively. 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  $(\beta_i(\mathbf{M}, \mathbf{b}_1, \dots, \mathbf{b}_m))_{i=1}^j$  or  $(\gamma_i(\mathbf{M}^T, \mathbf{c}_1, \dots, \mathbf{c}_p))_{i=1}^j$  is fully exhausted when  $m$  or  $p$  deflations have occurred, respectively. The iterative procedure is then stopped, giving equal numbers of left and right Lanczos vectors.

The non-zero elements of  $\mathbf{T}_n$  in Eq. (34) and of  $\tilde{\mathbf{T}}_n$  in Eq. (35), as determined by Algorithm 5.1 of Ref. [37], enforce the biorthogonality between the  $m + p + 1$  vectors that can overlap in exact arithmetic [35]. Our implementation is consistent with this algorithm, with the exception that the following procedure is added to the beginning of Step (1)

```

if  $n > 1$  then
  for  $k = 1$  to  $\max\{1, n - p_c - 1\}$  do
     $\hat{\mathbf{v}}_n \leftarrow \hat{\mathbf{v}}_n - \mathbf{v}_k(\mathbf{w}_k^T \hat{\mathbf{v}}_n)$ 
  end for
end if

```

and the following to the beginning of Step (2)

```

if  $n > 1$  then
  for  $k = 1$  to  $\max\{1, n - m_c - 1\}$  do
     $\hat{\mathbf{w}}_n \leftarrow \hat{\mathbf{w}}_n - (\hat{\mathbf{w}}_n^T \mathbf{v}_k)\mathbf{v}_k$ 
  end for
end if

```

These additions lead to an algorithm that enforces the biorthogonality between all Lanczos vectors in inexact arithmetic,

$$\mathbf{W}_n^T \mathbf{V}_n = \mathbf{\Delta}_n = \text{diag}(\delta_1, \delta_2, \dots, \delta_m), \quad (36)$$

and not just between the vectors that can overlap in exact arithmetic. We observe that this modification of the algorithm is important for the 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].

We stop the algorithm when it reaches a given chain length  $n$  (i.e. maximum number of iterations). The algorithm then generates, as Algorithm 5.1 of Ref. [37], the  $n \times n$  matrices

$$\begin{aligned} \mathbf{T}_n^{\text{P}} &= \mathbf{\Delta}_n^{-1} \mathbf{W}_n^T \mathbf{M} \mathbf{V}_n \\ &= \mathbf{T}_n + \mathbf{\Delta}_n^{-1} \mathbf{W}_n^T \mathbf{V}_n^{\text{D}}, \end{aligned} \quad (37)$$

$$\begin{aligned} \tilde{\mathbf{T}}_n^{\text{P}} &= (\mathbf{W}_n^T \mathbf{M} \mathbf{V}_n \mathbf{\Delta}_n^{-1})^T \\ &= \tilde{\mathbf{T}}_n + ((\mathbf{W}_n^{\text{D}})^T \mathbf{V}_n \mathbf{\Delta}_n^{-1})^T, \end{aligned} \quad (38)$$

related by

$$\mathbf{\Delta}_n \mathbf{T}_n^{\text{P}} = (\tilde{\mathbf{T}}_n^{\text{P}})^T \mathbf{\Delta}_n, \quad (39)$$

which are banded when no deflations have occurred [35].

The matrix  $\mathbf{T}_n^{\text{P}}$  can be viewed as the oblique projection of  $\mathbf{M}$  onto  $\mathcal{K}_j(\mathbf{M}, \mathbf{b}_1, \dots, \mathbf{b}_m)$  and orthogonally to  $\mathcal{K}_j(\mathbf{M}^T, \mathbf{c}_1, \dots, \mathbf{c}_p)$  [37]. Diagonalization of the matrix yields  $n$  eigenvalues, which approximate the eigenvalues of  $\mathbf{M}$ , and associated right and left eigenvectors. The right eigenvectors can be transformed to approximate eigenvectors of  $\mathbf{M}$  by premultiplication by  $\mathbf{V}_n$ , and the left eigenvectors to approximate eigenvectors of  $\mathbf{M}^T$  by premultiplication by  $\mathbf{W}_n^T \mathbf{\Delta}_n^{-1}$  [35]. Approximate eigenvectors with dominant (low- and high-lying) eigenvalues are typically better converged than the ones in the middle [39–41].

### 1. Choice of starting vectors

In Appendix A, we demonstrate how the expressions for matrix elements between excited states  $m$  and ground or excited initial states are linear in the excited determinant components  $r_{\mu m}$  and  $l_{m\mu}$ . This observation has previously been used to construct starting vectors based on excited states in EOM-CC theory [42], and have therein been used to justify the use of the starting vectors that we employ in the band Lanczos algorithm [36, 42, 43].

Starting vectors based on the ground state are given by

$$b_{\mu 0}^X = \xi_{\mu}^X, \quad (40)$$

$$c_{0\nu}^X = \text{EOM}\eta_{\nu}^X - X_{00}\bar{t}_{\nu}, \quad (41)$$

and starting vectors based on excited state  $m$  are given by [42]

$$b_{\mu m}^X = \sum_{\nu} \left( \text{EOM}A_{\mu\nu}^X + \delta_{\mu\nu} X_{00} - \xi_{\mu}^X \bar{t}_{\nu} \right) r_{\nu m}, \quad (42)$$

$$c_{m\nu}^X = \sum_{\mu} l_{m\mu} \left( \text{EOM}A_{\mu\nu}^X + \delta_{\mu\nu} X_{00} - \xi_{\mu}^X \bar{t}_{\nu} \right). \quad (43)$$



The specification of  $\xi_\mu^X$ ,  ${}^{\text{EOM}}\eta_\nu^X$ ,  $X_{00}$  and  ${}^{\text{EOM}}A_{\mu\nu}^X$  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 Lanczos vectors,

$$\begin{aligned} \mathbf{b}_i^X &= \sum_j \mathbf{v}_j (\mathbf{w}_j^T \mathbf{b}_i^X) \\ &= \sum_{j=1}^m \mathbf{v}_j b_{ji}^X, \end{aligned} \quad (44)$$

$$\begin{aligned} (\mathbf{c}_i^X)^T &= \sum_j \left( (\mathbf{c}_i^X)^T \mathbf{v}_j \right) \mathbf{w}_j^T \\ &= \sum_{j=1}^p c_{ij}^X \mathbf{w}_j^T. \end{aligned} \quad (45)$$

The sum in Eq. (44) is restricted since  $\mathbf{b}_i^X \in \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$  while each  $(\mathbf{w}_{m+1}, \dots)$  is biorthogonal to all  $(\mathbf{v}_1, \dots, \mathbf{v}_m)$ . Likewise, the sum in Eq. (45) is restricted since  $\mathbf{c}_i^X \in \text{span}\{\mathbf{w}_1, \dots, \mathbf{w}_p\}$  while each  $(\mathbf{v}_{p+1}, \dots)$  is biorthogonal to all  $(\mathbf{w}_1, \dots, \mathbf{w}_p)$ .

Thus, the transition moments to excited states can be obtained by contracting the starting vectors with the vectors of excited state  $n$ ,

$$\begin{aligned} X_{ni} &= \mathbf{L}_n^T \mathbf{b}_i^X \\ &= \sum_{j=1}^m (\mathbf{L}_n^T \mathbf{v}_j) b_{ji}^X \\ &= \sum_{j=1}^m L_{nj} b_{ji}^X, \end{aligned} \quad (46)$$

$$\begin{aligned} X_{in} &= (\mathbf{c}_i^X)^T \mathbf{R}_n \\ &= \sum_{j=1}^p c_{ij}^X (\mathbf{w}_j^T \mathbf{R}_n) \\ &= \sum_{j=1}^p c_{ij}^X R_{jn}. \end{aligned} \quad (47)$$

$R_{jn}$  and  $L_{nj}$  are simply the components of the right and left eigenvectors of  $\mathbf{T}_n^{\text{P}}$ , respectively, and  $b_{ji}^X = \mathbf{w}_j^T \mathbf{b}_i^X$  and  $c_{ij}^X = \mathbf{c}_i^X \mathbf{v}_j$  are products of starting vectors and biorthonormalized Lanczos vectors.

### E. Generation of a reduced basis

The iterative process that is used to calculate sets of excited EOM-CC states is given below. We take the set  $J_0$  to contain the indices of already calculated EOM-CC states, and start from the ground state, giving  $J_0 = \{0\}$ . The iterative procedure for the  $c$ -th EOM-CC state calculation is as follows.

1. Choose a chain length,  $n_c$ , and a subset of the state indices from previous calculations,  $I_c \subseteq J_{c-1}$ . Also

choose a set of operators  $X_c$  based on the final states accessed by the operators (see Section IV).

2. Sequences of right and left start vectors,  $(\mathbf{b}_i^X)_{i \in I_c, X \in X_c}$  and  $(\mathbf{c}_i^X)_{i \in I_c, X \in X_c}$ , are constructed in accordance with Eqs. (40) to (43).
3. The band Lanczos algorithm described in Section IID is run with the chain length  $k_c$ , the field-free Jacobian matrix  $\mathbf{A}^{(0)}$ , and the sequences of starting vectors, constructing the matrix  $\mathbf{T}_{n_c}^{\text{P}}$ .
4. After the band Lanczos algorithm has finished, the eigenvalues and corresponding right and left eigenvectors of  $\mathbf{T}_{n_c}^{\text{P}}$  are calculated. Together, these determine a set of approximate EOM-CC states indexed by  $\{m\}$ .
5. States  $m$  with eigenvalues  $\omega_m^c > \omega^{\text{max}}$  are discarded.
6. Matrix elements for all operators  $X \in X_c$  and combinations of final  $m$  and initial states  $i \in I_c$  are calculated in accordance with Eq. (47) and Eq. (46). States  $m$  with  $X_{im} X_{mi} = S_{im}^X < S^{\text{min}}$  for all operators and initial states are discarded.
7. Remaining right and left eigenvectors  $m$  of  $\mathbf{T}_{n_c}^{\text{P}}$  are transformed to approximate right and left eigenvectors  $\mathbf{R}_m$  and  $\mathbf{L}_m$  of  $\mathbf{A}^{(0)}$  by premultiplication by  $\mathbf{V}_n$  and  $\mathbf{W}_n^T \mathbf{\Delta}_n^{-1}$ , respectively. States with vectors linearly independent of previously calculated vectors are stored. The indices of the stored states are added to the previous index set  $J_c = J_{c-1} \cup \{m\}$ .

The iterative procedure is repeated if states of higher excitation levels are desired. Afterwards, the excited-state Jacobian and overlap matrices, with elements  $\tilde{A}_{mn}^{(0)} = \mathbf{L}_m^T \mathbf{A}^{(0)} \mathbf{R}_n$  and  $\tilde{S}_{mn} = \mathbf{L}_m^T \mathbf{R}_n$ , 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  $\tilde{\mathbf{A}}^{(0)}$ . 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 Eq. (25) and Eq. (28).

## III. COMPUTATIONAL DETAILS

Experimental geometries from the NIST database [44] are used for LiH, LiF and C<sub>2</sub>H<sub>4</sub>. 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<sub>s</sub>

symmetry for the chosen geometry, with the  $xy$  plane as the mirror plane.

In all following calculations, the aug-cc-pCVDZ basis set [45, 46] is used for atoms targeted by the core-exciting pulses; the aug-cc-pVDZ [45] basis set is adopted for the remaining atoms in the molecules. Valence and core states are obtained with the asymmetric band Lanczos algorithm with varying chain lengths as specified in Section IV.

Lanczos vectors with Euclidian norms of less than  $1 \times 10^{-9}$  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  $1 \times 10^{-7}$  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  $n_\gamma$ -photon transition to the desired excited states. The carrier frequency  $\omega_i$  and the frequency RMS width  $\sigma_i^\omega = 1/(2\sigma_i^t)$  are parameters of the pulse providing photon  $i$ , see Section II A.

Core-valence separation (CVS) [47–49] 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. 1, 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  $1 \times 10^{-2}$  for valence states and  $1 \times 10^{-1}$  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 Section II E.

A fixed pump-probe delay of 40 a.u. (about 0.967 552 fs) is used for both 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 0 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 [50] with a maximum time step of 0.1 a.u., and maximum and minimum local errors of  $1 \times 10^{-7}$  and  $1 \times 10^{-9}$ , respectively, see Appendix B.

Each component of the time-dependent dipole moment expectation value and electric field vectors are multiplied with the Hann window before Fourier transformation.

All calculations are performed using a development version of the  $e^T$  program [51].

## IV. RESULTS AND DISCUSSION

### A. Lithium fluoride: convergence and non-linear 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 0.005 a.u., is in Fig. 1 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 [28], where the F  $K$ -edge is targeted by the probe pulse. 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. 1 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 converged right or left vector residual norms will be discarded in the following band Lanczos calculations.

In Fig. 2, 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 10 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

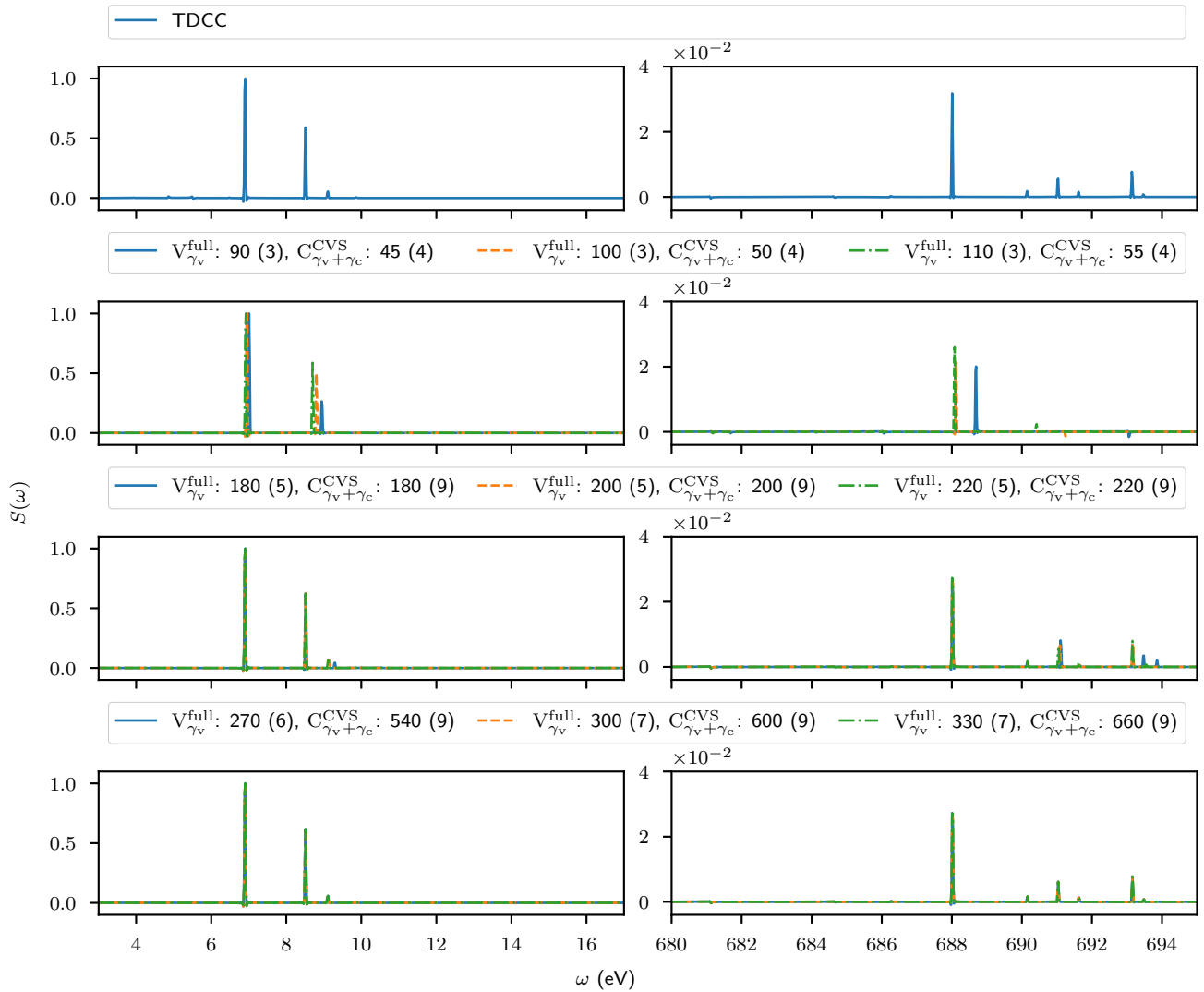


FIG. 1. LiF pump-probe absorption  $S(\omega)$  as a function of frequency  $\omega$  in the valence and core regions, normalized by the tallest peaks in the spectra. The TDCC results are shown in the top left and right panels. TD-EOM-CC results, calculated at different band Lanczos chain lengths, are 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  $\gamma_v$  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).

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. 2, 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 prac-

tically 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 non-linear interactions are involved. The embedded Dormand-Prince 5(4) integrator is seen to perform well for TD-EOM-CC.

The bottom panel of Fig. 2 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 non-



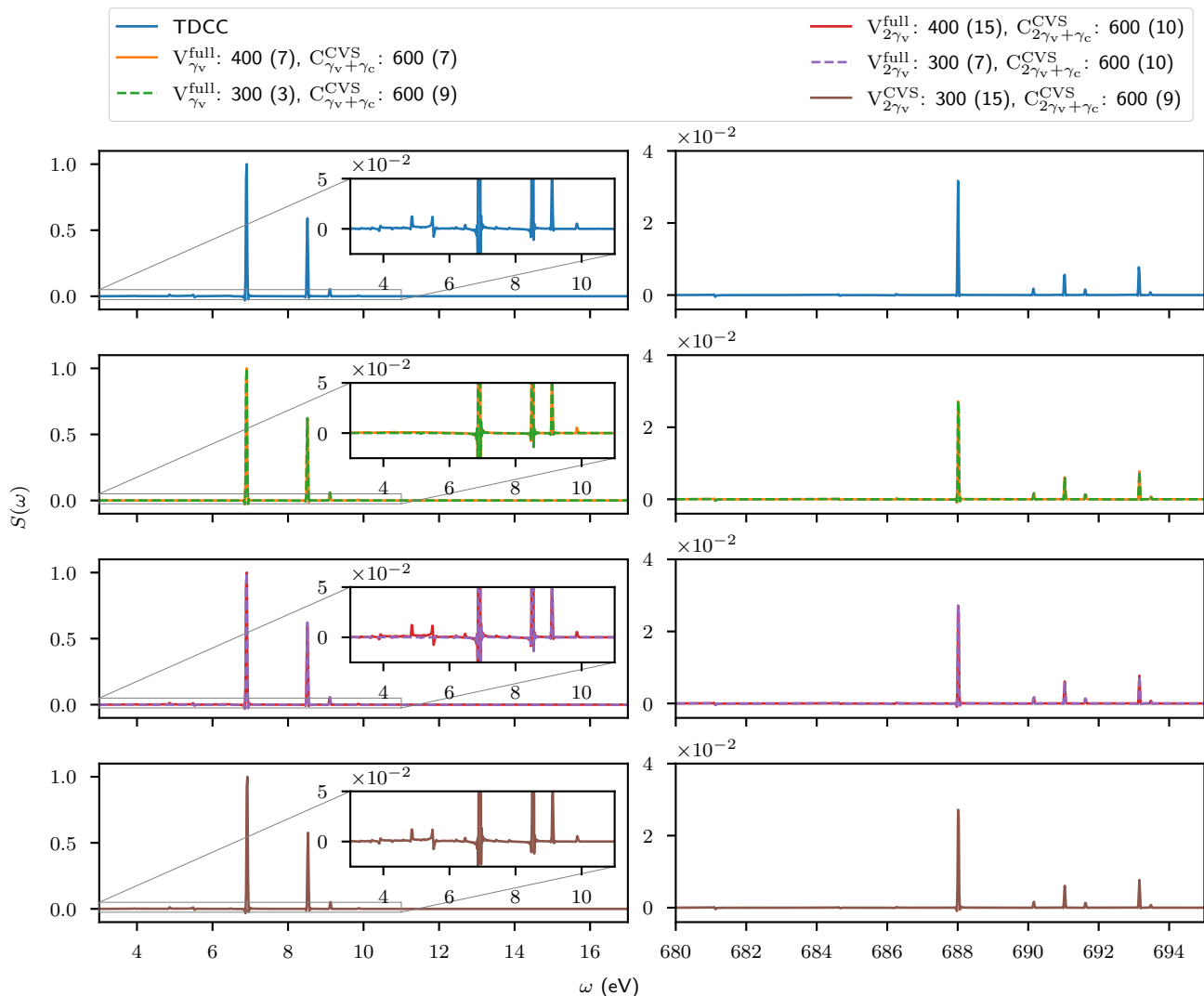


FIG. 2. LiF pump-probe absorption  $S(\omega)$  as a function of frequency  $\omega$  in the valence and core regions, normalized by the tallest peaks in the spectra. The TDCC results are shown in the top left and right panels. TD-EOM-CC results, calculated at different band Lanczos chain lengths, are 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,  $\gamma_v$  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).

projected 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]. 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. 3, where the latter is taken from Ref. [28]. In all core state calculations a fixed band Lanczos chain length of 400 is used. However, the number of converged core states, given in brackets, differs due to

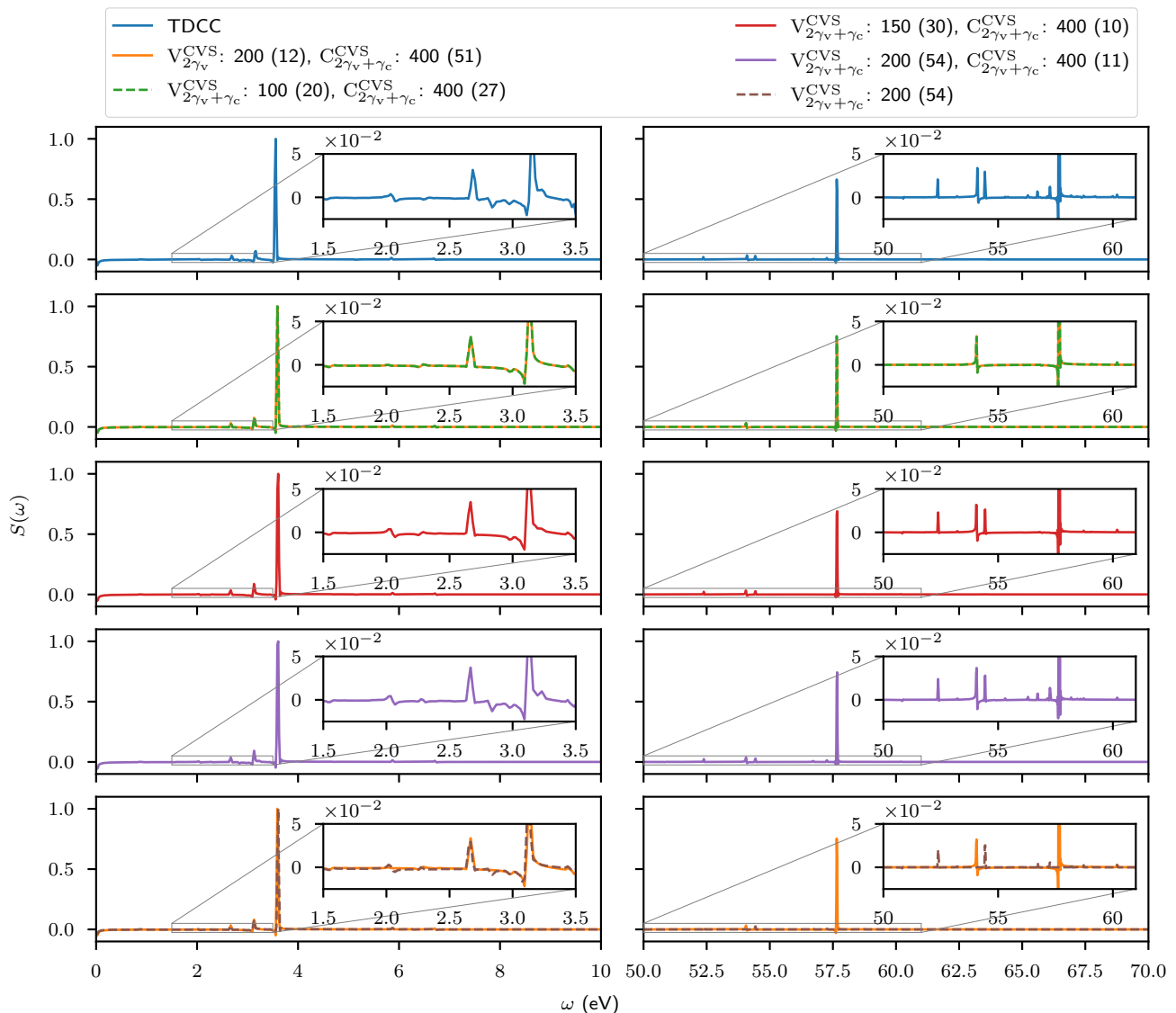


FIG. 3. LiH pump-probe absorption  $S(\omega)$  as a function of frequency  $\omega$  in the valence and core regions, normalized by the tallest peaks in the spectra. The TDCC results are shown in the top left and right panels. TD-EOM-CC results, calculated at different band Lanczos chain lengths, are 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 different starting vectors employed.

Since TD-EOM-CC with energy-limited valence and core states successfully reproduced the TDCC spectrum of LiF in Section 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. 3. 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, as 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. 3. The valence chain lengths used are 100 (green), 150 (red) and 200 (purple). The number of converged states in the valence region in-

creases 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.

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. 3. 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 *per se* 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.

### C. Ethylene: non-linear 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 FWHM durations of about 403 as and 201 as, respectively. The carrier frequency of the pump pulse is set to 8.003 337 eV, and the probe pulse to 285.608 346 eV (C  $K$ -edge). 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. 4, is characterised 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 Section 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. 4, 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 can indicate that more secondary valence excited states should be in-

cluded in the computation. It may also be a signature of fundamental differences in two-photon absorption as described by TD-EOM-CC and TDCC.

### D. Glycine: transient absorption

As a final example, we use computational procedure to model attosecond transient absorption by the glycine molecule. The RMS widths of  $(0.490072x + 0.871682y)$ -polarized pump and  $z$ -polarized 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 6.358 895 eV, and the probe pulse to 405.305 168 eV (N  $K$ -edge). 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 start vectors and a chain length of 1500, gives 17 converged states. A subsequent core state calculation, with ground and valence state start vectors and a chain length of 3000, gives 20 converged states.

As a note of caution, the maximum energy of the valence states is 10.458 724 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, meaning that smaller features may be missing in the spectra.

In the top left panel of Fig. 5, 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 fs to 2.902 656 fs in intervals of 60.472 as. 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 normal-

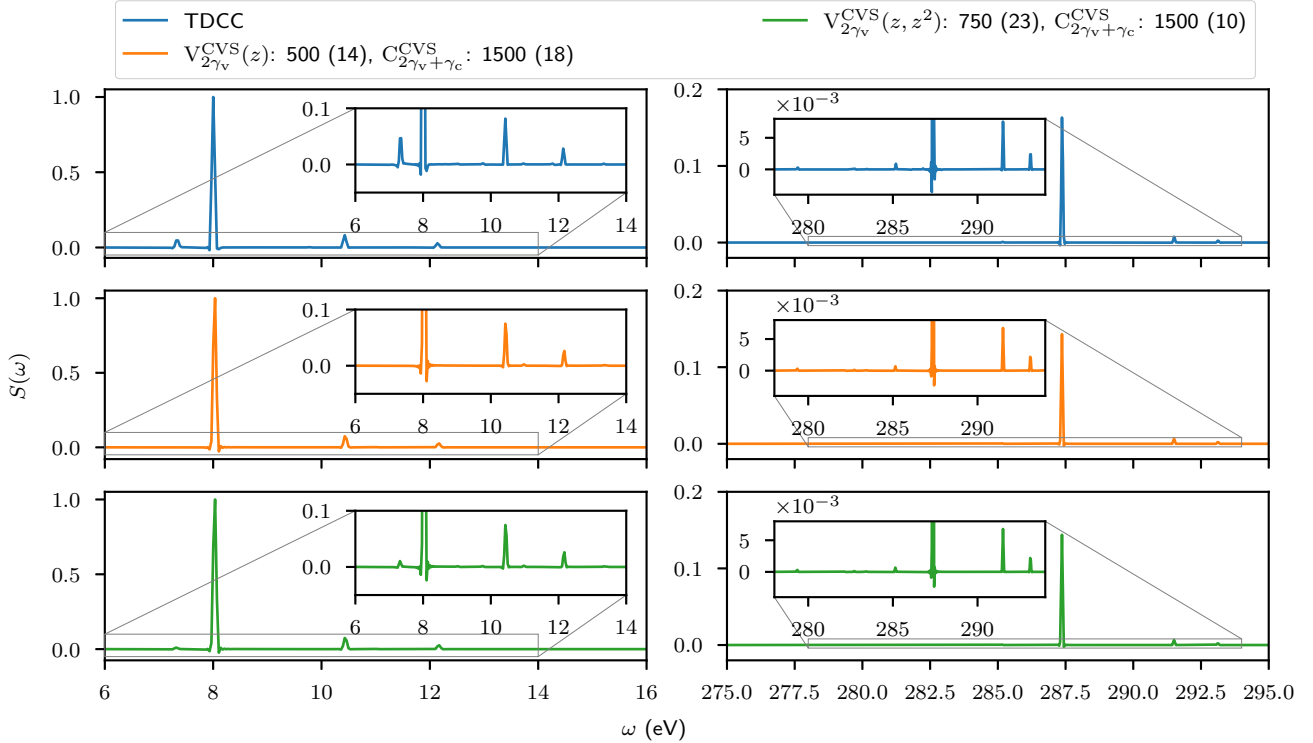


FIG. 4. Ethylene pump-probe absorption  $S(\omega)$  as a function of frequency  $\omega$  in the valence and core regions, normalized by the tallest peaks in the spectra. The TDCC results are shown in the top left and right panels, and TD-EOM-CC results in the lower panels. In the middle panels, a ground 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).

izing by the tallest peak in the ground-state probe spectrum. In the bottom left panel, the difference spectra at 1.269 912 fs and 1.572 272 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. 5, 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. 5, 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. 5, 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 [52, 53].

## 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 ac-

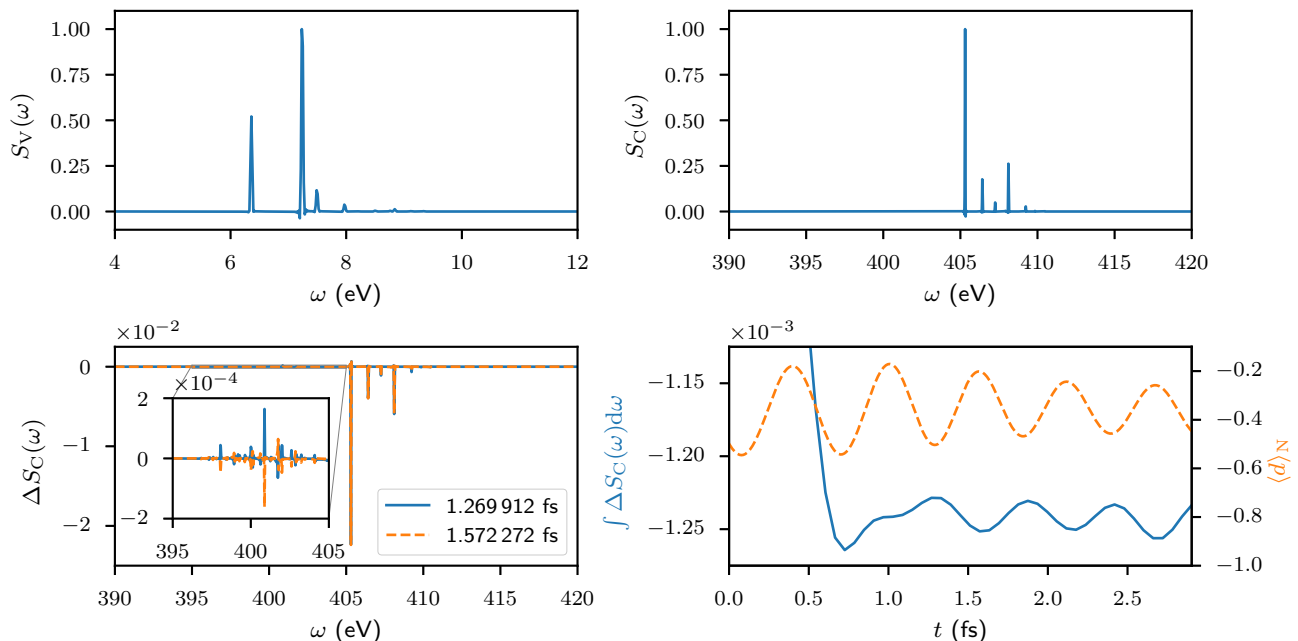


FIG. 5. Glycine. Top left: pump pulse absorption from the ground state  $S_V(\omega)$ , normalized by the tallest peak in the spectrum. Top right: probe pulse absorption from the ground state  $S_C(\omega)$ , normalized by the tallest peak in the spectrum. Bottom left: pump-probe absorption minus the 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.

count. The specific start vectors used in the calculations direct the band Lanczos algorithm towards states that are useful for representing the interactions. The start 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 Section 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 Section 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 Section 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 Section 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.

## ACKNOWLEDGMENTS

We acknowledge Sarai D. Folkestad for her role in implementing the simple Lanczos algorithm and density-based excited-state transition moments, available in the  $e^T$  program, which serve as precursors to the implementations here. We acknowledge financial support from The Research Council of Norway through FRINATEK Project Nos. 263110 and 275506, from the Marie Skłodowska-Curie European Training Network *COmputational Spectroscopy In Natural sciences and Engineering (COSINE)*, Grant Agreement No. 765739, and from the Independent Research Fund Denmark, DFF-FNU Research Project 2 No. 7014-00258B. Computing resources through UNINETT Sigma2—the National Infrastructure for High Performance Computing and Data Storage in



Norway (Project No. NN2962k) and through the Center for High Performance Computing (CHPC) at SNS are also acknowledged. Finally, the COST Action CA18222, *AttoChem*, is acknowledged.

## AUTHOR'S CONTRIBUTIONS

S.C. and H.K. conceptualized and supervised the project. T.M., S.C. and A.S.S. implemented and optimized the band Lanczos solver. A.B. and A.S.S. did the initial TDCC implementation, and A.S.S. the TD-EOM-CC extension. A.C.P. implemented procedures for calculating eigenstates and matrix elements of combined sets of EOM-CC states. A.S.S. and T.M. carried out the calculations and interpreted the results. All authors contributed to the writing of the manuscript.

### 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

$$\begin{aligned}
X_{ij} &= \langle \tilde{\psi}_i | X | \psi_j \rangle \\
&= \sum_{\kappa\lambda} l_{i\kappa} \langle \kappa | \bar{X} | \lambda \rangle r_{\lambda j} \\
&= \sum_{\nu} l_{i0} \langle \text{HF} | \bar{X} | \nu \rangle r_{\nu j} + \sum_{\mu\nu} l_{i\mu} \langle \mu | \bar{X} | \nu \rangle r_{\nu j} \\
&\quad + \left( l_{i0} X_{00} + \sum_{\mu} l_{i\mu} \xi_{\mu}^X \right) r_{0j} \\
&= \sum_{\nu} l_{i0} \langle \text{HF} | [\bar{X}, \tau_{\nu}] | \text{HF} \rangle r_{\nu j} \\
&\quad + \sum_{\mu\nu} l_{i\mu} \left( \text{LR} A_{\mu\nu}^X + \langle \mu | \tau_{\nu} \bar{X} | \text{HF} \rangle \right) r_{\nu j} \\
&\quad + \left( l_{i0} X_{00} + \sum_{\mu} l_{i\mu} \xi_{\mu}^X \right) r_{0j} ,
\end{aligned} \tag{A1}$$

where

$$\text{LR} A_{\mu\nu}^X = \langle \mu | [\bar{X}, \tau_{\nu}] | \text{HF} \rangle , \tag{A2}$$

$$X_{00} = \langle \text{HF} | \bar{X} | \text{HF} \rangle , \tag{A3}$$

$$\xi_{\mu}^X = \langle \mu | \bar{X} | \text{HF} \rangle . \tag{A4}$$

#### 1. Left excited and right ground state

The left excited state  $m$  has the reference component  $l_{m0} = 0$  (cf. Eq. (23)), while the right ground state has components  $r_{00} = 1$ ,  $r_{\mu 0} = 0$  (cf. 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 [43]. Note that this expression is linear in the excited determinant components  $l_{m\mu}$ .

#### 2. Left ground and right excited state

The left ground state has the components  $l_{00} = 1$ ,  $l_{0\mu} = \bar{t}_{\mu}$  (cf. Eq. (19)), while the right excited state  $m$  has the reference component  $r_{0m} = -\sum_{\mu} \bar{t}_{\mu} r_{\mu m}$  (cf. Eq. (22)). Inserting this into Eq. (A1), we obtain

$$\begin{aligned}
X_{0n} &= \sum_{\nu} \left( \text{LR} \eta_{\nu} + \sum_{\mu} \bar{t}_{\mu} \langle \mu | \tau_{\nu} \bar{X} | \text{HF} \rangle \right) r_{\nu n} \\
&\quad - \left( X_{00} + \sum_{\mu} \bar{t}_{\mu} \xi_{\mu}^X \right) \sum_{\nu} \bar{t}_{\nu} r_{\nu n} ,
\end{aligned} \tag{A6}$$

where

$$\text{LR} \eta_{\nu} = \langle \text{HF} | [\bar{X}, \tau_{\nu}] | \text{HF} \rangle + \sum_{\mu} \bar{t}_{\mu} \text{LR} A_{\mu\nu}^X . \tag{A7}$$

The term  $\sum_{\nu} \text{LR} \eta_{\nu} r_{\nu n}$  appears in CC response theory [43], and the other terms are specific to EOM-CC. Equation (A6) is equivalent to Eq. (65) in [54] and can also be written as

$$X_{0n} = \sum_{\nu} \left( \text{EOM} \eta_{\nu}^X - X_{00} \bar{t}_{\nu} \right) r_{\nu n} , \tag{A8}$$

where

$$\begin{aligned}
\text{EOM} \eta_{\nu}^X &= \text{LR} \eta_{\nu}^X + \sum_{\mu} \bar{t}_{\mu} \langle \mu | \tau_{\nu} \bar{X} | \text{HF} \rangle \\
&\quad - \left( \sum_{\mu} \bar{t}_{\mu} \xi_{\mu}^X \right) \bar{t}_{\nu}
\end{aligned} \tag{A9}$$

cf. Eq. (18) of Ref. [55]. Note that Eq. (A8) is linear in the excited determinant components  $r_{\nu n}$ .

#### 3. Left and right excited states

The left excited state  $m$  has the reference component  $l_{m0} = 0$ , while the right excited state  $n$  has  $r_{0n} = -\sum_{\mu} \bar{t}_{\mu} r_{\mu n}$  (cf. Eqs. (22) to (23)). Inserting this into Eq. (A1), we obtain

$$\begin{aligned}
X_{mn} &= \sum_{\mu\nu} l_{m\mu} \left( \text{LR} A_{\mu\nu}^X + \langle \mu | \tau_{\nu} \bar{X} | \text{HF} \rangle \right) r_{\nu n} \\
&\quad - \sum_{\mu} l_{m\mu} \xi_{\mu}^X \sum_{\nu} \bar{t}_{\nu} r_{\nu n} \\
&= \sum_{\mu\nu} l_{m\mu} \left( \text{LR} A_{\mu\nu}^X + \langle \mu | \tau_{\nu} \bar{X} | \text{HF} \rangle - \xi_{\mu}^X \bar{t}_{\nu} \right) r_{\nu n} \\
&= \sum_{\mu\nu} l_{m\mu} \left( \text{EOM} A_{\mu\nu}^X + \delta_{\mu\nu} X_{00} - \xi_{\mu}^X \bar{t}_{\nu} \right) r_{\nu n} ,
\end{aligned} \tag{A10}$$

where

$$\begin{aligned} \text{EOM}A_{\mu\nu}^X &= \langle \mu | \bar{X} | \nu \rangle - \delta_{\mu\nu} X_{00} \\ &= \text{LR}A_{\mu\nu}^X + \langle \mu | \tau_\nu \bar{X} | \text{HF} \rangle - \delta_{\mu\nu} X_{00} \end{aligned} \quad (\text{A11})$$

cf. Eq. (20) in Ref. [55].

The term  $\sum_{\mu\nu} l_{m\mu} \text{LR}A_{\mu\nu}^X r_{\nu n}$  appears in CC response theory [42], and the other terms are specific to EOM-CC. 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) [50]. This method yields both fourth- and fifth-order accurate solutions at each time

step, and is specified by the Butcher tableau [50, 56]

$$\begin{array}{c|cccccc} 0 & & & & & & \\ \frac{1}{5} & \frac{1}{5} & & & & & \\ \frac{3}{10} & \frac{3}{40} & \frac{9}{40} & & & & \\ \frac{4}{5} & \frac{44}{45} & -\frac{56}{15} & \frac{32}{9} & & & \\ \frac{8}{9} & \frac{19372}{6561} & -\frac{25360}{2187} & \frac{64448}{6561} & -\frac{212}{729} & & \\ 1 & \frac{9017}{3168} & -\frac{355}{33} & \frac{46732}{5247} & \frac{49}{176} & -\frac{5103}{18656} & \\ 1 & \frac{35}{384} & 0 & \frac{500}{1113} & \frac{125}{192} & -\frac{2187}{6784} & \frac{11}{84} \\ \hline & \frac{35}{384} & 0 & \frac{500}{1113} & \frac{125}{192} & -\frac{2187}{6784} & \frac{11}{84} & 0 \\ & \frac{5179}{57600} & 0 & \frac{7571}{16695} & \frac{393}{640} & -\frac{92097}{339200} & \frac{187}{2100} & \frac{1}{40} \end{array} \quad (\text{B1})$$

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,

$$\epsilon_{\mathcal{O}(4)} = \left\| \mathbf{y}_{\mathcal{O}(5)} - \mathbf{y}_{\mathcal{O}(4)} \right\|_2. \quad (\text{B2})$$

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 Section 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 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* **530**, 66 (2016).  
 [2] M. Galli, V. Wanie, D. P. Lopes, E. P. Månsson, A. Trabattini, 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. Photonics* **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. Glowina, 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. Photonics* **14**, 30 (2020).
- [5] A. Malyzhenkov, Y. P. Arbelo, P. Craievich, P. Dijkstra, E. Ferrari, S. Reiche, T. Schietinger, P. Juranić, and E. Prat, *Phys. Rev. Res.* **2**, 042018 (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. Bodenstern, 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: At. Mol. Opt. Phys.* **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, *Non-Hermitian eigenvalue problems*, in *Templates for the Solution of Algebraic Eigenvalue Problems* (SIAM, 2000) Chap. 7, pp. 149–231.
- [36] I. H. Godtlielsen and O. Christiansen, *Phys. Chem. Chem. Phys.* **15**, 10035 (2013).
- [37] R. W. Freund, *Acta Numer.* **12**, 267–319 (2003).
- [38] N. Shimizu, T. Mizusaki, Y. Utsuno, and Y. Tsunoda, *Comp. Phys. Commun.* **244**, 372 (2019).
- [39] J. K. Cullum and R. A. Willoughby, *Lanczos Algorithms For Large Symmetric Eigenvalue Computations, Vol. 1: Theory* (IBM T. J. Watson Research Center, 1984).
- [40] G. H. Golub and C. F. V. Loan, *Matrix Computations 4th Edition* (The Johns Hopkins University Press, 2013).
- [41] Y. Saad, *Iterative Methods for Sparse Linear Systems* (SIAM, 2003).
- [42] B. N. C. Tenorio, M. A. C. Nascimento, A. B. Rocha, and S. Coriani, *J. Chem. Phys.* **151**, 184106 (2019).
- [43] S. Coriani, T. Fransson, O. Christiansen, and P. Norman, *J. Chem. Theory Comput.* **8**, 1616 (2012).
- [44] NIST Computational Chemistry Comparison and Benchmark Database, NIST Standard Reference Database Number 101 (Release 20 August 2019).
- [45] R. A. Kendall, T. H. Dunning, and R. J. Harrison, *J. Chem. Phys.* **96**, 6796 (1992).
- [46] D. E. Woon and T. H. Dunning, *J. Chem. Phys.* **103**, 4572 (1995).
- [47] L. S. Cederbaum, W. Domcke, and J. Schirmer, *Phys. Rev. A* **22**, 206 (1980).
- [48] S. Coriani and H. Koch, *J. Chem. Phys.* **143**, 181103 (2015).
- [49] B. N. C. Tenorio, T. Moitra, M. A. C. Nascimento, A. B. Rocha, and S. Coriani, *J. Chem. Phys.* **150**, 224104 (2019).
- [50] J. Dormand and P. Prince, *J. Comput. Appl. Math.* **6**, 19 (1980).
- [51] 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).
- [52] A. D. Dutoi, K. Gokhberg, and L. S. Cederbaum, *Phys. Rev. A* **88**, 013419 (2013).
- [53] A. D. Dutoi and L. S. Cederbaum, *Phys. Rev. A* **90**, 023414 (2014).
- [54] H. Koch, R. Kobayashi, A. Sanchez de Merás, and P. Jørgensen, *J. Chem. Phys.* **100**, 4393 (1994).
- [55] R. Faber and S. Coriani, *J. Chem. Theory Comput.* **15**, 520 (2019).
- [56] A. Iserles, Error control, in *A First Course in the Numerical Analysis of Differential Equations*, Cambridge Texts in Applied Mathematics (Cambridge University Press, 2008) p. 105–122, 2nd ed.