

New and Efficient Equation-of-Motion Coupled-Cluster Framework for Core-Excited and Core-Ionized States

Marta L. Vidal,[†] Xintian Feng,^{‡,§} Evgeny Epifanovsky,[§] Anna I. Krylov,^{||} and Sonia Coriani^{*,†}

[†]DTU Chemistry, Department of Chemistry, Technical University of Denmark, Kongens Lyngby DK-2800, Denmark

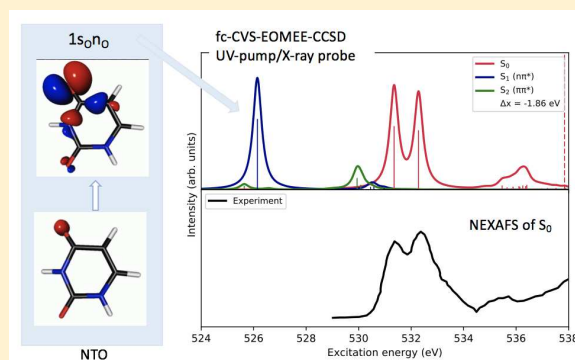
[‡]Department of Chemistry, University of California, Berkeley, California 94720, United States

[§]Q-Chem Incorporated, 6601 Owens Drive, Suite 105, Pleasanton, California 94588, United States

^{||}Department of Chemistry, University of Southern California, Los Angeles, California 90089-0482, United States

Supporting Information

ABSTRACT: We present a fully analytical implementation of the core–valence separation (CVS) scheme for the equation-of-motion (EOM) coupled-cluster singles and doubles (CCSD) method for calculations of core-level states. Inspired by the CVS idea as originally formulated by Cederbaum, Domcke, and Schirmer, pure valence excitations are excluded from the EOM target space and the frozen-core approximation is imposed on the reference-state amplitudes and multipliers. This yields an efficient, robust, practical, and numerically balanced EOM-CCSD framework for calculations of excitation and ionization energies as well as state and transition properties (e.g., spectral intensities, natural transition, and Dyson orbitals) from both the ground and excited states. The errors in absolute excitation/ionization energies relative to the experimental reference data are on the order of 0.2–3.0 eV, depending on the K-edge considered and on the basis set used, and the shifts are systematic for each edge. Compared to a previously proposed CVS scheme where CVS was applied as a posteriori projection only during the solution of the EOM eigenvalue equations, the new scheme is computationally cheaper. It also achieves better cancellation of errors, yielding similar spectral profiles but with absolute core excitation and ionization energies that are systematically closer to the corresponding experimental data. Among the presented results are calculations of transient-state X-ray absorption spectra, relevant for interpretation of UV-pump/X-ray probe experiments.



1. INTRODUCTION

By providing tunable high-energy radiation, advanced light sources, such as X-ray free electron laser (X-FEL) and synchrotron installations, enable a variety of X-ray-based spectroscopies.^{1–3} Recent advances in beam quality greatly expanded possible applications of X-rays, giving rise to a proliferation of techniques, including those operating in time-resolved and nonlinear regimes.^{1–5} Fundamentally, these spectroscopies exploit electronic transitions involving core orbitals. Since core-level binding energies are characteristic of a species, X-ray absorption and X-ray photoelectron spectroscopies (XAS and XPS, respectively) are powerful techniques for probing the electronic structure of atoms and molecules.⁶ The localized nature of core orbitals makes X-ray-based spectroscopies sensitive to local environment. However, just as in the case of VUV-based techniques,⁷ theoretical modeling is required to unambiguously assign spectral features and to relate experimental measurements to molecular structures. The experimental advances over the past decade have been driving the interest in developing highly accurate theoretical methods for X-ray spectroscopy and, in particular, for time-resolved

XAS, which is proving to be a powerful means to investigate molecular dynamics.^{5,8–10}

Owing to its low computational cost relative to *ab initio* methods, time-dependent (TD) density functional theory (DFT) has been among the most commonly used techniques for modeling absorption spectra, including X-ray absorption.¹¹ However, TDDFT often fails to deliver an accurate description of spectroscopic properties. TDDFT core excitations computed with conventional exchange–correlation functionals, for instance, are grossly underestimated with MAE of the order of 20 eV,¹¹ and time-independent DFT-based procedures like Δ KS, transition-potential DFT (TP-DFT),^{12,13} and orthogonality constrained DFT (OC-DFT)^{14,15} are found to perform better than TDDFT for X-ray absorption.

Thus, there is an evident need for reliable wave function-based methods for calculations of XAS spectra of both ground and excited states. Although more expensive, these methods can be systematically improved, thereby yielding to results of controlled accuracy.¹⁶ The focus of this contribution is on

Received: January 16, 2019

Published: April 9, 2019



extending coupled-cluster (CC) approaches,^{17–24} which are among the most successful *ab initio* methods for molecular properties and electronic spectra in the UV–vis region, to absorption and ionization phenomena in the X-ray domain.^{8,25–28}

Although ionization or excitations of core electrons superficially appear to be similar to valence transitions, the numerical experiments have shown that direct application of standard approaches to core-level transitions leads to unsatisfactory results,⁵ due to the following essential features of core-ionized and core-excited states. First, these states have open-shell character.²⁹ Second, the core-level states lie very high in energy (hundreds of electron volts, depending on the edge). Third, orbital relaxation effects are much more important for core states than for valence states because the outer orbitals are more delocalized and better shielded from the nuclear charge than the tight and localized core orbitals. Fourth, these transitions appear to be much more sensitive to the effect of the environment.^{30,31} Fifth, these high-lying states are metastable with respect to electron ejection,^{32,33} i.e., they are Feshbach resonances that can autoionize via two-electron transitions in which one valence electron fills the core hole and a second valence electron is ejected. Thus, they are embedded in the ionization continuum, and their description within Hermitian quantum mechanics is problematic.

Multistate methods, such as equation-of-motion coupled cluster (EOM-CC),^{17,18,20,22–24,34} can effectively describe multiconfigurational wave functions and tackle open-shell character and orbital relaxation effects in one computational scheme. The effect of the environment can be included by using a variety of implicit or explicit solvent models. However, dealing with high-energy interior states, especially with those embedded in autodetaching continuum, is more problematic. Although the Davidson procedure can be modified to solve for the eigenstates dominated by the desired transition (MOM-like³⁵) or lying within the desired energy range,^{26,36} the convergence and numerical stability are strongly affected by the presence of the continuum. An attempt to compute such states often produces pseudocontinuum states in which one electron occupies the most diffuse orbital.³⁷ Because in Hermitian quantum mechanics the resonances are not represented by a single state but rather by an increased density of states in the continuum,³⁸ it is not possible to obtain converged results by systematically improving the basis set—resonances simply dissolve in the continuum. In moderate bases, resonances appear as isolated states, but their representation in a discretized continuum is inherently prone to numeric instabilities. Although the EOM-CC methods can describe resonances by using complex-scaled and CAP-augmented approaches,³³ such calculations are much more expensive than regular bound-state calculations. Here we focus on alternative, less expensive approximate methods for modeling the spectroscopic properties of the core-level states.

Our strategy for tackling issues due to the continuum nature of the core levels is based on employing the core–valence separation (CVS) approximation proposed by Cederbaum et al. in 1980.³⁹ In essence, within the CVS scheme the continuum is projected out, such that core-level states become artificially stabilized (bound). By decoupling the core excitations from the rest of the configurational space, the CVS allows one to extend standard methods for excited and ionized states to the core-level states.^{28,40–42} This separation of full space into the bound part and the continuum is

reminiscent of the Feshbach–Fano treatment of resonances.^{43,44} Effectively, CVS results in a decoupling of the highly excited core states from the continuum of valence excitations; it also leads to a significant reduction of the computational costs. The core states can also be decoupled from the continuum by excluding double excitations from the configurational space and treating them perturbatively.³⁰ Although numerically stable, this approach suffers from insufficient description of orbital relaxation.³⁰ Because the CVS scheme admits selected double excitations that are crucial for describing orbital relaxation, the CVS-EOM-CCSD (or, similarly, the CVS variant of linear-response CCSD, LR-CCSD)²⁸ ansatz retains sufficient flexibility, resulting in accuracy comparable to that of regular EOM-CCSD/LR-CCSD,^{25,27,45} for those (small) cases where full EOM-CCSD or LR-CCSD calculations are feasible.²⁸

The analogy with the Feshbach–Fano formalism^{43,44} makes it clear that the CVS ansätze are *not* approximations to the full parent ansätze—rather, they should be thought of as the diabaticized versions of the full methods. For example, CVS-EOM-CCSD is not an approximation to EOM-CCSD but a diabaticized version of it in which the continuum part of the spectra is projected out. Consequently, hierarchical inclusion of higher excitations would ultimately converge to CVS-FCI (full configuration interaction), which is a diabaticized representation of the FCI solution. One can recover the exact position of the core states (and also their lifetimes) by solving the Schrödinger equation with the non-Hermitian effective Hamiltonian⁴³ obtained using the Löwdin projection formalism,⁴⁶ but this is beyond the standard CVS scheme, which has been shown to yield sufficiently reliable estimates of energies of core-level states within this diabatic picture. Just as in the practical implementations of the Feshbach–Fano approach,⁴⁷ one needs to define projectors separating the bound part of the spectra from the continuum. This step involves certain arbitrariness: it disappears in the exact limit,^{43,46} but the results for the diabaticized states depend on the choice of the projector. As explained below, the present version of CVS-EOM-CCSD employs a slightly different projector than the original one.²⁸

An implementation of the CVS within EOM-CC and LR-CC theories has been reported by Coriani and Koch.²⁸ In their work, the CVS was deployed as an a posteriori projection applied at each iteration of the solution of the EOM/LR-CC equations to eliminate excitations that do not involve at least one core electron, whereas the ground-state amplitudes and Lagrangian multipliers retained all possible types of excitations. Here, a different strategy is presented: the zero-order wave function parameters are computed within the frozen-core (fc) approximation and the subsequent EOM/LR-CC equations are solved imposing the core–valence separation analytically. The respective state and transition properties such as oscillator strengths, natural transition orbitals (NTOs), Dyson orbitals, exciton descriptors, etc., are obtained from the appropriate densities between different target-state manifolds. Specifically, transition properties can be computed between the ground state and a core-excited state as well as between a valence-excited and a core-excited state. This yields an efficient scheme for simulating, e.g., near-edge absorption fine structure (NEXAFS) spectra of both ground and excited states. The latter is required to simulate UV-pump/X-ray-probe experiments and to obtain core-ionization potentials and Dyson orbitals for XPS spectra of medium-size molecules. We report

illustrative results for NEXAFS and core IEs of all K-edges in neon, water, ammonia, ethylene, vinylfluoride, ozone, and adenine as well as the transient (time-resolved) NEXAFS spectrum of uracil.

2. THEORY

2.1. EOM-CCSD. In the EOM-CC approach, the target-state wave functions are parametrized using the following ansatz^{18,20,22,48}

$$\begin{aligned} \langle \Psi_L | &= \langle \Phi_0 | L e^{-T} \\ | \Psi_R \rangle &= e^T R | \Phi_0 \rangle \end{aligned} \quad (1)$$

where $|\Phi_0\rangle$ is a reference Slater determinant and T , R , and $L^\dagger$ are excitation operators.

The excitation operator T is the cluster operator⁴⁹

$$\begin{aligned} T &= \sum_{\mu} t_{\mu} \tau_{\mu} = T_1 + T_2 + \dots + T_N \\ T_1 &= \sum_{ia} t_i^a a_a^\dagger a_i, \quad T_2 = \frac{1}{4} \sum_{ijab} t_{ij}^{ab} a_a^\dagger a_b^\dagger a_j a_i, \quad \dots \end{aligned} \quad (2)$$

with τ_{μ} being the excitation operator and t_{μ} the corresponding cluster amplitudes determined by the CC equations for the reference state

$$\langle \Phi_{\mu} | \bar{H} - E_{CC} | \Phi_0 \rangle = 0; \quad E_{CC} = \langle \Phi_0 | \bar{H} | \Phi_0 \rangle \quad (3)$$

where $\langle \Phi_{\mu} |$'s represent μ -tuple excited determinants and $\bar{H}$ is the similarity transformed Hamiltonian

$$\bar{H} = e^{-T} H e^T \quad (4)$$

The operator R is a general excitation operator

$$R = R_0 + R_1 + \dots + R_n \equiv \sum_{\mu} r_{\mu} \tau_{\mu} \quad (5)$$

and L is a de-excitation operator

$$L = L_0 + L_1 + \dots + L_n \equiv \sum_{\mu} l_{\mu} \tau_{\mu}^\dagger \quad (6)$$

The choice of the reference state and the exact form of R and L depend on the EOM method to be used.^{20,22} Furthermore, in practical calculations the excitation and de-excitation operators must be truncated to some tractable level of excitation. In this work, we focus on the EOM-CCSD family of methods in which the cluster operator T is truncated after single (S) and double (D) excitations and so are the excitation R and de-excitation L operators.

One of the most obvious applications of EOM theory is the calculation of electronically excited states.^{18,48} In this approach, referred to as EOM-EE (EOM for excitation energies), the optimal reference state is usually the closed-shell ground-state Hartree–Fock determinant and the R and L operators conserve the number of electrons and their spin taking the following forms

$$R^{EE} = r_0 + \sum_{ia} r_i^a a_a^\dagger a_i + \frac{1}{4} \sum_{ijab} r_{ij}^{ab} a_a^\dagger a_b^\dagger a_j a_i + \dots \quad (7)$$

$$L^{EE} = \sum_{ia} l_i^a a_i^\dagger a_a + \frac{1}{4} \sum_{ijab} l_{ij}^{ab} a_i^\dagger a_j^\dagger a_b a_a + \dots \quad (8)$$

Another common use of EOM is calculations of ionization energies by the EOM-IP (EOM for ionization potentials) method.^{50–52} In this case, the excitation operator changes the number of electrons in the system, assuming the following form

$$R^{IP} = \sum_i r_i a_i + \frac{1}{4} \sum_{ij a} r_{ij}^a a_a^\dagger a_j a_i + \dots \quad (9)$$

The EOM amplitudes r_{μ} and l_{μ} are found as stationary points of the EOM functional

$$E = \frac{\langle \Psi_L | H | \Psi_R \rangle}{\langle \Psi_L | \Psi_R \rangle} \quad (10)$$

By applying the bivariational principle,^{53,54} one arrives at the nonsymmetric eigenvalue problem

$$\langle \Phi_{\mu} | \bar{H} - E | R \Phi_0 \rangle = 0; \quad \langle \Phi_0 | L | \bar{H} - E | \Phi_{\mu} \rangle = 0 \quad (11)$$

where the eigenvectors of the Hamiltonian form a biorthogonal set¹⁸

$$\langle \Phi_0 | L_i | R_j | \Phi_0 \rangle = \delta_{ij} \quad (12)$$

Thus, implementation of the EOM-CCSD method boils down to diagonalization of the effective Hamiltonian $\bar{H}$ in the basis of the reference and singly and doubly excited determinants, which can be written in matrix form as

$$\bar{H} = \begin{pmatrix} E_{CC} & \bar{H}_{OS} & \bar{H}_{OD} \\ 0 & \bar{H}_{SS} & \bar{H}_{SD} \\ 0 & \bar{H}_{DS} & \bar{H}_{DD} \end{pmatrix} \quad (13)$$

giving rise to the EOM-CCSD right and left eigenvalue equations

$$\begin{pmatrix} \bar{H}_{SS} - E_{CC} & \bar{H}_{SD} \\ \bar{H}_{DS} & \bar{H}_{DD} - E_{CC} \end{pmatrix} \begin{pmatrix} R_1 \\ R_2 \end{pmatrix} = \omega \begin{pmatrix} R_1 \\ R_2 \end{pmatrix} \quad (14)$$

$$\begin{pmatrix} L_1 & L_2 \end{pmatrix} \begin{pmatrix} \bar{H}_{SS} - E_{CC} & \bar{H}_{SD} \\ \bar{H}_{DS} & \bar{H}_{DD} - E_{CC} \end{pmatrix} = \omega \begin{pmatrix} L_1 & L_2 \end{pmatrix} \quad (15)$$

where ω is the energy difference with respect to the reference state.

In practice, eqs 14 and 15 are solved iteratively using a generalized Davidson's iterative diagonalization procedure^{55–57} which requires computation of the right σ and left $\tilde{\sigma}$ trial vectors

$$\sigma_1 = ([\bar{H}_{SS} - E_{CC}] R_1)_1 + (\bar{H}_{SD} R_2)_1 \quad (16)$$

$$\sigma_2 = (\bar{H}_{DS} R_1)_2 + ([\bar{H}_{DD} - E_{CC}] R_2)_2 \quad (17)$$

$$\tilde{\sigma}_1 = (L_1 [\bar{H}_{SS} - E_{CC}])_1 + (L_2 \bar{H}_{DS})_1 \quad (18)$$

$$\tilde{\sigma}_2 = (L_1 \bar{H}_{SD})_2 + ([L_2 \bar{H}_{DD} - E_{CC}])_2 \quad (19)$$

In EOM-CC, molecular properties are often calculated as expectation values^{18,54} using unrelaxed one-particle density matrices like

$$\gamma_{pq} = \langle \Psi_L | a_p^\dagger a_q | \Psi_R \rangle \quad (20)$$

In the present study, we focus on transition properties involving core states, specifically, the oscillator strengths f

between the ground state and the core-excited states and between valence-excited and core-excited states. These can be formulated via contractions of property integral matrices and transition density matrices between state i and j , for instance

$$f(i \rightarrow j) = \frac{2}{3}(\omega_j - \omega_i) \sum_{\alpha=x,y,z} \mu_{\alpha}^{i \rightarrow j} \mu_{\alpha}^{j \rightarrow i} \quad (21)$$

where

$$\mu_{\alpha}^{i \rightarrow j} = \sum_{pq} \mu_{pq}^{\alpha} \gamma_{pq}^{i \rightarrow j}; \quad \mu_{\alpha}^{j \rightarrow i} = \sum_{rs} \mu_{rs}^{\alpha} \gamma_{rs}^{j \rightarrow i} \quad (22)$$

Above, μ^{α} refers to a specific Cartesian component of the electric dipole operator and the state i can be either the ground state or an excited state.

The transition density matrices $\gamma_{pq}^{i \rightarrow j}$ and $\gamma_{pq}^{j \rightarrow i}$ are generalizations of eq 20, and their specific form in the fc-CVS-EOM-CCSD framework is discussed in the next sections. Here we used the so-called unrelaxed EOM transition densities, which differ from the density matrices derived within the LR-CC formalism in that they do not contain terms originating from the amplitudes' response to the external perturbation. In the framework of EOM-CC, the amplitude-response terms can also be included by using, for example, a Lagrangian approach, giving rise to the relaxed state and transition density matrices, which are identical to those in the LR-CC formalism.^{17,50,58–61} Furthermore, orbital relaxation to the external perturbation is also omitted, as typically done within coupled-cluster response theory.¹⁹

Transition density matrices describe the changes in electron density upon excitation and can be interpreted as exciton wave function^{62–65} when expressed in coordinate space

$$\chi(r_e, r_h) = \sum_{pq} \gamma_{pq} \phi_p(r_e) \phi_q(r_h) \quad (23)$$

where r_e and r_h denote particle (electron) and hole coordinates, respectively (using $r_e = r_p = r$, χ is reduced to the transition density). Properties of the exciton can be quantified by various expectation values, i.e., exciton size, hole–particle separation, and correlation, etc. These exciton descriptors help to assign state characters.^{64–66} Applying a singular value decomposition (SVD) to the transition matrices yields the NTOs,^{64–69} which can be used to describe the electronic excitations in terms of hole–particle excitons

$$\chi(r_e, r_h) = \sum_K \sigma_K \tilde{\phi}_K^e(r_e) \tilde{\phi}_K^h(r_h) \quad (24)$$

where $\tilde{\phi}_K^e$ and $\tilde{\phi}_K^h$ are particle and hole orbitals obtained by SVD of γ and σ_K are the corresponding singular values. Usually only a few σ_K are nonzero, giving rise to simple interpretation of excited-state characters in terms of one-electron excitations. Note that the sum of σ_K^2 is equal to the squared norm of γ , which provides a simple metric quantifying the one-electron character of the transition, i.e., for pure single excitations, $\|\gamma\| = 1$.

2.2. fc-CVS-EOM-CCSD Method. The essence of CVS is the separation between the bound and the continuum parts of the spectrum, which allows one to compute the bound part of core-level states. The separation depends on the choice of a projector, which is not uniquely defined. This arbitrariness is akin to the inherent arbitrariness of diabatic representation of the electronic Hamiltonian: the diabatic states are not uniquely defined, but if the full problem is solved exactly then the final

answer does not depend on a specific choice of diabatic basis. Different flavors of CVS can be defined by different choices of the projectors, and one way to assess the relative merit of a particular choice is by judging the numeric performance of the resulting methods with respect to a chosen reference, either theoretical or experimental. In discussing diabaticization schemes, for instance, the magnitude of the coupling elements is often used as a guideline, so better choices of diabatic states yield smaller couplings. The discussion about various CVS schemes has been commonly framed in the literature in terms of the analysis of the matrix elements of the Hamiltonian that couple core-level and valence states and/or core orbitals and valence orbitals. In the discussion below, we follow the second line of reasoning and analyze the difference between different CVS schemes in terms of the Coulomb integrals that are set to zero.

Different variants of separation between core and valence excitations have been realized for a variety of approximate wave function methods as well as for TDDFT, and we refer the interested reader to ref 5 for a recent overview. As mentioned in the [Introduction](#), within the CVS-CC framework of ref 28, an effective core–valence separation was introduced (as a projection) in the solution of the CC eigenvalue equations since, physically, the core-excited states are energetically far away from the valence excited states.

Already in 1980, Cederbaum et al.³⁹ suggested that since core orbitals are strongly localized in space and energetically are well separated from the valence orbitals one can decouple core and valence orbitals by zeroing out the respective blocks of the model Hamiltonian, such that the solutions of the Schrödinger equation can be separated into the core and valence domains. In 1985, Barth and Schirmer⁴⁰ implemented this idea within the second-order algebraic diagrammatic construction method ADC(2)^{70,71} by observing that, in view of negligible coupling between core and valence orbitals, all of the following two-electron Coulomb integrals are negligible

$$\begin{aligned} \langle I|lqr\rangle &= \langle p|lqr\rangle = \langle pq|lr\rangle = \langle pqlr\rangle \approx 0 \\ \langle I|lpq\rangle &= \langle pq|IJ\rangle \approx 0 \\ \langle IJ|Kp\rangle &= \langle IJ|pK\rangle = \langle Ip|JK\rangle = \langle pI|JK\rangle \approx 0 \end{aligned} \quad (25)$$

where small letters indicate general valence orbitals and capital letters core orbitals. As a consequence, one can omit the blocks of the ADC propagator matrix $\mathbf{M}$ involving these integrals and therefore decouple pure valence excitations from excitations involving core electrons. This idea was later extended to all other members of the ADC family of propagator methods.^{41,71–73} Since MP2 is an underlying wave function for the ADC propagator,⁷¹ omitting the integrals in eq 25 immediately corresponds to using the frozen-core approximation for the MP2 ground-state energy.

If the above criterion of negligible Coulomb integrals is applied within the CCSD ansatz in the canonical basis,⁷⁴ the resulting CCSD energy and the amplitude equations for the “valence-only” amplitudes t_{iv}^a and t_{ijv}^{ab} are identical to those obtained within the frozen-core approximation, that is, assuming that all ground-state amplitudes and multipliers where at least one occupied index refers to a core orbital are zero.

Inspired by this observation, we here propose to use the frozen core approximation during determination of the ground-state amplitudes t_{μ} and Lagrangian multipliers λ_{μ}

while constraining the excitation/ionization operators R and L to involve at least one core orbital, which introduces restrictions on the left and right EOM equations and on the respective density matrices. We name the resulting approach as the fc-CVS-EOM-CCSD method.

The programmable expressions for obtaining the right and left excitation vectors within the fc-CVS-EOM-CCSD methods have been derived from ref 54 and for the density matrices from ref 60 and can be found in the [Supporting Information](#).

As an illustrative example, we show below how the expression of the right linearly transformed vector elements σ_i^a is modified for the fc-CVS case. The general expression for σ_i^a is⁵⁴

$$\begin{aligned} \sigma_i^a = & \sum_b F_{ab} r_i^b - \sum_j F_{ij} r_j^a - \sum_{jb} I_{ibja}^1 r_j^b + \sum_{jb} F_{jb} r_{ij}^{ab} \\ & - \frac{1}{2} \sum_{jkb} I_{jki}^6 r_{jk}^{ab} - \frac{1}{2} \sum_{jbc} I_{jab}^7 r_{ij}^{bc} \end{aligned} \quad (26)$$

with

$$\begin{aligned} F_{ia} = & f_{ia} + \sum_{jb} t_j^b \langle ij || ab \rangle \\ F_{ij} = & f_{ij} + \sum_a t_i^a f_{ja} + \sum_{ka} t_k^a \langle jk || ia \rangle + \sum_{kab} t_i^a t_k^b \langle jk || ab \rangle \\ & + \frac{1}{2} \sum_{kbc} t_{jk}^{ac} \langle jk || bc \rangle \\ F_{ab} = & f_{ab} - \sum_i t_i^a f_{ja} - \sum_{ic} t_i^c \langle ia || bc \rangle + \sum_{ijc} t_i^c t_j^a \langle ij || bc \rangle \\ & - \frac{1}{2} \sum_{jkc} t_{jk}^{ac} \langle jk || bc \rangle \\ I_{iajb}^1 = & \langle ia || jb \rangle - \sum_k t_k^b \langle jk || ia \rangle - \sum_c t_i^c \langle jc || ac \rangle \\ & + \sum_{kc} t_i^c t_k^b \langle jk || ac \rangle - \sum_{kc} t_{ik}^{bc} \langle jk || ac \rangle \\ I_{ijka}^6 = & \langle ij || ka \rangle - \sum_c t_k^c \langle ij || ac \rangle \\ I_{iabc}^7 = & \langle ia || bc \rangle - \sum_j t_j^a \langle ij || bc \rangle \end{aligned} \quad (27)$$

It is convenient to further split the occupied orbitals ($i, j, k, l, \dots$) into the two sub-blocks: occupied valence labeled with an additional v subindex, ($i_v, j_v, k_v, l_v, \dots$) and occupied core orbitals denoted by a capital letter ($I, J, K, L, \dots$). Our fc-CVS scheme then entails (i) reducing the set of occupied orbitals to only the core ones in the excitation process and (ii) freezing the core orbitals in optimization of the ground state CC wave function parameters (amplitudes and multipliers). Hence, by restricting the EOM excitations to core excitations only, all terms involving only valence excitations disappear in the equations above. The frozen-core approximation further simplifies the equations since the terms in eq 27 involving ground-state amplitudes vanish for the core orbitals or, in other words, only ground-state amplitudes with valence-occupied orbitals are retained in the fc-CVS case.

The fc-CVS expression for the linearly transformed trial vector in eq 26 thus reads as follows

$$\begin{aligned} \sigma_I^a = & \sum_b F_{ab} r_I^b - \sum_J F_{IJ} r_J^a - \sum_{Jb} I_{IbJa}^1 r_J^b + \sum_{Jb} F_{Jb} r_{IJ}^{ab} \\ & + \sum_{j_v b} F_{j_v b} r_{Ij_v}^{ab} - \sum_{Jk_v b} \langle Jk_v || Ib \rangle r_{Jk_v}^{ab} - \frac{1}{2} \sum_{JKb} \langle JK || Ib \rangle r_{JK}^{ab} \\ & - \frac{1}{2} \sum_{Jbc} I_{Jabc}^7 r_{IJ}^{bc} - \frac{1}{2} \sum_{j_v bc} I_{j_v abc}^7 r_{Ij_v}^{bc} \end{aligned} \quad (28)$$

with

$$\begin{aligned} F_{Ia} = & f_{Ia} + \sum_{j_v b} t_{j_v}^b \langle Ij_v || ab \rangle \\ F_{IJ} = & f_{IJ} + \sum_{k_v a} t_{k_v}^a \langle Jk_v || Ia \rangle \\ F_{i_v a} = & f_{i_v a} + \sum_{j_v b} t_{j_v}^b \langle i_v j_v || ab \rangle \\ I_{IaJb}^1 = & \langle Ia || Jb \rangle - \sum_{k_v} t_{k_v}^b \langle Jk_v || Ia \rangle \\ I_{Iabc}^7 = & \langle Ia || bc \rangle - \sum_{j_v} t_{j_v}^a \langle Ij_v || bc \rangle \end{aligned} \quad (29)$$

We followed the same strategy to derive the expressions for other linearly transformed vector blocks and for the (transition) density matrices; all programmable expressions are given in the [Supporting Information](#).

We emphasize that we only used the integral screening as per eq 25 to justify the use of the frozen-core approximation in the determination of the ground-state CC wave function parameters and to write the effective Hamiltonian as given in eq 13. The linearly transformed vectors used to determine the target state energies and the transition properties are derived only invoking the frozen-core condition on the ground-state parameters while restricting the EOM excitations to core excitations (the latter referred to as the CVS condition²⁸). Therefore, some terms in the intermediates still contain such integrals. Although small, these terms are found to further slightly improve the flexibility of our ansatz to address, via correlation, the relaxation effects that follow core excitation. Strictly speaking, the complete neglect of the integrals in eq 25 defines yet another variant of a CVS scheme, which will not be discussed in the present study.

We conclude this section by discussing the size extensivity of our fc-CVS-EOM-CCSD model. As explained, for instance, in refs 75 and 76 the size extensivity of the total energy and size intensity of the excitation energies is guaranteed if the singles and doubles block of the first column of the matrix representation of the effective Hamiltonian $\bar{H}$ are zero, as per eq 13. The blocks in question contain the amplitude-constraint terms

$$\Omega_\mu = \langle \Phi_\mu | \exp(-T) H \exp(T) | \Phi_0 \rangle \quad (30)$$

where for CCSD $\langle \Phi_\mu |$ is either a single- or a double-excitation determinant with respect to the reference Slater determinant $|\Phi_0\rangle$, see eq 3.

Obviously, in fc-CVS-EOM-CCSD

$$\Omega_{i_v}^a = \langle \Phi_{i_v}^a | \exp(-T) H \exp(T) | \Phi_0 \rangle = 0 \quad (31)$$

$$\Omega_{i_v j_v}^{ab} = \langle \Phi_{i_v j_v}^{ab} | \exp(-T) H \exp(T) | \Phi_0 \rangle = 0 \quad (32)$$

since these terms are identical to the valence-only amplitude equations obtained by application of the frozen-core approximation. As for the remaining terms in which the occupied indices are either a core I , a core J , or both, let us consider, for instance, the elements Ω_I^a . In the frozen-core approximation, they become

$$\begin{aligned}\Omega_I^a &= \langle \Phi_I^a | \exp(-T) H \exp(T) | \Phi_0 \rangle \\ &= f_{aI} - \sum_{k_v} f_{k_v I} t_{k_v}^a + \sum_{k_v} \langle k_v a || cI \rangle t_{k_v}^c - \frac{1}{2} \sum_{k_v l_v} \langle k_v l_v || cI \rangle t_{k_v l_v}^{ca} \\ &\quad - \sum_{k_v l_v} \langle k_v l_v || cI \rangle t_{k_v}^c t_{l_v}^a\end{aligned}\quad (33)$$

In the canonical basis, the respective Fock-matrix terms are zero. If one invokes the negligible integrals conditions as per eq 25, these blocks would become exactly zero. Similar arguments apply for the blocks $\Omega_{I'v}^{ab}$ and Ω_{IJ}^{ab} . We note in passing that one can alternatively consider the Ω_I^a , $\Omega_{I'v}^{ab}$, and Ω_{IJ}^{ab} blocks when only the negligible integral conditions are applied, set them equal to zero, and solve for the amplitudes t_I^a , $t_{I'v}^{ab}$, and t_{IJ}^{ab} . It is then straightforward to prove that the frozen-core conditions $t_I^a = 0$, $t_{I'v}^{ab} = 0$, and $t_{IJ}^{ab} = 0$ are trivial solutions to the equations $\Omega_I^a = 0$, $\Omega_{I'v}^{ab} = 0$, and $\Omega_{IJ}^{ab} = 0$.

However, since we defined our method by only invoking the frozen-core approximation during the optimization of the ground state, we ensure that the fc-CVS-EOM-CCSD approach remains size extensive (i.e., size intensive for the core excitation energies) by setting per construction the above-mentioned blocks of $\bar{H}$ to zero. The “negligible integrals condition” suggests in any case that they will be small and can therefore be neglected with some degree of confidence.

Numerical evidence of size intensivity is shown in Table S3 of the SI, where we report core-excitation energies computed for a cluster of Ne and H₂O separated by $R = 100$ bohr, as well as for each system alone. Exactly the same core-excitation energies are obtained in both cases.

3. COMPUTATIONAL DETAILS

We implemented the fc-CVS-EOM-CC method in the Q-Chem electronic structure package^{77,78} using the libtensor library.⁷⁹ The geometries of H₂O, NH₃, CO, C₂H₄, C₂H₃F, and O₃ were optimized at the fc-CCSD(T)/cc-pVQZ level using CFOUR.⁸⁰ For 9H-adenine, we considered both planar and nonplanar structures taken from the literature. The nonplanar structure was optimized at the fc-RI-MP2/cc-pVTZ level,⁸¹ whereas the planar structure was optimized at the B3LYP/cc-pVTZ level.⁸² In the TR-NEXAFS simulations of uracil we used different structures: an optimized MP2/cc-pVTZ ground-state structure, a ground-state structure and an S₁ minimum both optimized at the SF-BH&HLYP/6-31+G** level from ref 83, and two stationary-point (i.e., zero-gradient) structures for the S₁ and S₂ excited states obtained at the EOM-CCSD/aug-cc-pVDZ level of theory. All structures were planar or almost planar.

In the (TR-)NEXAFS calculations we considered several different basis sets, and results are here reported for Pople's 6-311++G** (pure d functions) and Dunning's aug-cc-pVTZ and aug-cc-pCVTZ sets. In selected cases, the basis sets were further augmented with uncontracted Rydberg-type functions whose exponents were computed according to the prescription of Kaufmann et al.⁸⁴ Using such system-specific Rydberg

functions affords a more compact description of Rydberg states than an alternative brute-force strategy of adding additional diffuse sets with even-tempered exponents.⁸⁵

To analyze the EOM states (i.e., the extent of Rydberg character), we considered, for a few selected cases, the spatial extent of the respective wave functions as well as the size of the particle NTO of a transition.⁶⁹ Both approaches deliver similar information; however, when using the former, one needs to consider the difference between the expectation value of the second moment of charges $\langle r^2 \rangle$ of the target EOM state and the CCSD reference⁸⁵ because the size of the electronic wave function depends on the system size.

Experimental data were taken from ref 86 for adenine, ref 87 for neon, ref 88 for H₂O and NH₃, refs 89 and 90 for CO, ref 91 for C₂H₄ and C₂H₃F, and ref 92 for O₃. The experimental NEXAFS spectrum of uracil is from ref 93. All experimental spectra were digitized from the original references using WebPlotDigitizer.⁹⁴ The spectra were generated using a Python script, and NTOs were visualized using MOLDEN.⁹⁵ The CVS-LR-CCSD results we compare with were obtained using a development version of Dalton.⁹⁶

4. RESULTS AND DISCUSSION

4.1. Near-Edge Absorption Spectroscopy and Core Ionization Energies. To test the performance of our method, we first considered the neon atom. Table S4 reports the computed excitation energies, oscillator strengths, and IEs with three different basis sets, all supplemented with Rydberg-type functions (with $n = 2.5$ –4.5); the corresponding spectra are shown in Figure 1. As the atomic NEXAFS spectra are due to transitions from the 1s orbital to Rydberg states,⁶ the inclusion of Rydberg-type functions (or, alternative, a large set of diffuse functions) is mandatory in order to reproduce the Rydberg

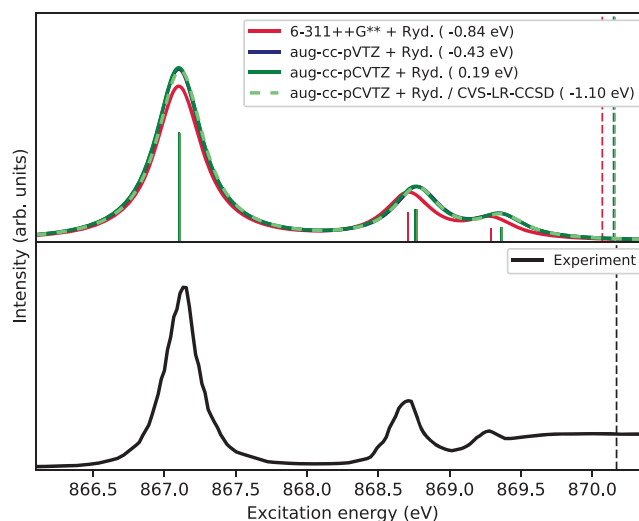


Figure 1. Neon. fc-CVS-EOM-CCSD X-ray absorption spectra obtained by convolution of the computed excitation energies and oscillator strengths with a Lorentzian function (fwhm = 0.4 eV). Experimental spectrum was digitized from ref 87. Vertical dashed lines correspond to the core ionization energies. Experimental IE is 870.17 eV. Energy shifts required to align the NEXAFS profiles in each basis set with the experimental one are indicated in parentheses. Computed IEs have been shifted by the same amount as used to align the NEXAFS profiles. NEXAFS (shifted) spectrum and IE obtained using the CVS-LR-CCSD method²⁸ for one basis set are also shown.

progression of peaks in the experimental spectrum.²⁷ The extent of Rydberg character in neon is quantified in Table S5. Table S4 and Figure 1 also show the results obtained with the CVS-LR-CCSD method in the largest of the three selected basis sets.

The NEXAFS spectra have been shifted, along with the IEs, to align with the first peak of the experimental NEXAFS spectrum (estimated to be at 867.10 eV). After the shift, the computed peaks match the experimental ones almost perfectly. Of the three sets, Dunning's aug-cc-pCVTZ(+Rydberg) yields the smallest absolute shift from experiment (+0.19 eV), followed by −0.43 eV of the aug-cc-pVTZ(+Rydberg) set, versus −0.84 eV of Pople's 6-311+G**(+Rydberg). The spectral profiles in the two Dunning sets overlap completely after the alignment.

With respect to previously reported CCSD results,²⁷ obtained in the aug-cc-pCVTZ basis supplemented with Rydberg functions using the Lanczos algorithm with all electrons correlated (i.e., no CVS), the absolute shift from experiment in the fc-CVS-EOM-CCSD method is lower and has an opposite sign (+0.19 versus −1.1 eV). The shift is also smaller than the one obtained using the CVS-LR-CCSD approach of ref 28, whereas the spectral profiles are practically the same, as it can be appreciated from results in both Table S4 and Figure 1. We also draw the reader's attention to Table S2, where the fc-CVS-EOMEE-CCSD results for neon, as well as for a few molecular systems, are compared with the corresponding CVS-LR-CCSD results taken from ref 28.

The NEXAFS and IE values of H₂O are reported in Table S6, with the corresponding spectra shown in Figure 2. The upper panel of Figure 2 shows the spectra for the chosen basis sets without Rydberg-type functions, whereas the two middle panels show those obtained including the Rydberg-type functions. The third panel in particular compares the spectra and IEs yielded by our fc-EOM-CCSD method and the CVS-LR-CCSD method. Besides an overall shift (taking the value

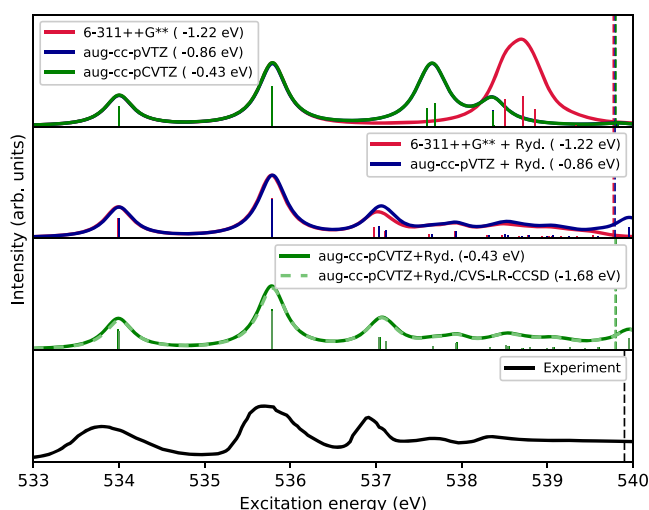


Figure 2. Water. fc-CVS-EOMEE-CCSD O K-edge X-ray absorption spectra obtained by convolution of the spectral data in Table S6 with a Lorentzian function (fwhm = 0.4 eV). Experimental spectrum (0.12 eV resolution) was digitized from ref 88. Dashed vertical lines correspond to the IEs. Energy shifts required to align the NEXAFS profiles in each basis set with the experimental one are indicated in parentheses. Computed IEs have been shifted by the same amount as used to align the NEXAFS profiles.

534.0 eV as reference for the experimental first peak maximum, which varies slightly for the three bases), the separation between the two first peaks is practically the same, whereas huge differences are observed for the other bands known to have Rydberg character. Both the relative intensity and the position of the third band and the following ones are strongly overestimated in the bases without Rydberg functions. Table S5 shows the sizes of particle NTOs and the $\langle r^2 \rangle$ values (i.e., the second moments of charge density), clearly revealing Rydberg character of certain transitions.

Also for H₂O the energy shifts required to realign with the experimental spectrum are smaller than those obtained using CVS-LR-CCSD,²⁸ see the third panel of Figure 2, as well as Table S7 and Table S2 in the SI. The spectral profiles, on the other hand, are basically identical.

Remarkably, in the aug-cc-pCVTZ basis the shift is smaller than in the aug-cc-pVTZ basis, whereas the reverse trend has been observed using the CVS-LR approach of ref 28. Thus, the current approach shows a systematic improvement (in terms of deviation from experiment) of the results with respect to the basis set increase.

Another system whose gas-phase NEXAFS is dominated by Rydberg states is NH₃. Figure 3 shows the computed spectra;

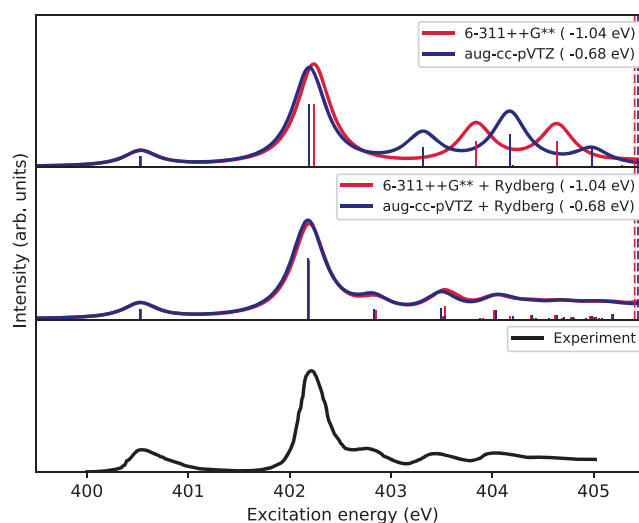


Figure 3. Ammonia. fc-CVS-EOMEE-CCSD N K-edge X-ray absorption spectra obtained from convolution of the spectral data in Table S8 with a Lorentzian function (fwhm = 0.4 eV). Experimental spectrum (0.2 eV resolution) was digitized from ref 88. Dashed vertical lines indicate the IEs. Energy shifts required to align the NEXAFS profiles in each basis set with the experimental one are indicated in parentheses. Computed IEs have been shifted by the same amount as used to align the NEXAFS profiles.

the raw data are in Table S8. The spectra were aligned with respect to the peak maximum of the first experimental band, estimated at 400.53 eV. As for the previous systems, the Dunning basis shows a smaller shift compared to the experimental peaks (−0.68 vs −1.04 eV). Neither Pople's 6-311+G** nor Dunning's aug-cc-pVTZ can correctly reproduce the third and higher bands without the inclusion of additional diffuse functions. The Rydberg character of these bands is clearly revealed by the data in Table S5. As in the previous two cases, the overall shift of the fc-CVS-EOMEE-CCSD spectra from the experimental one is smaller than that obtained with the CVS-LR-CCSD scheme,²⁸ see also Table S2.

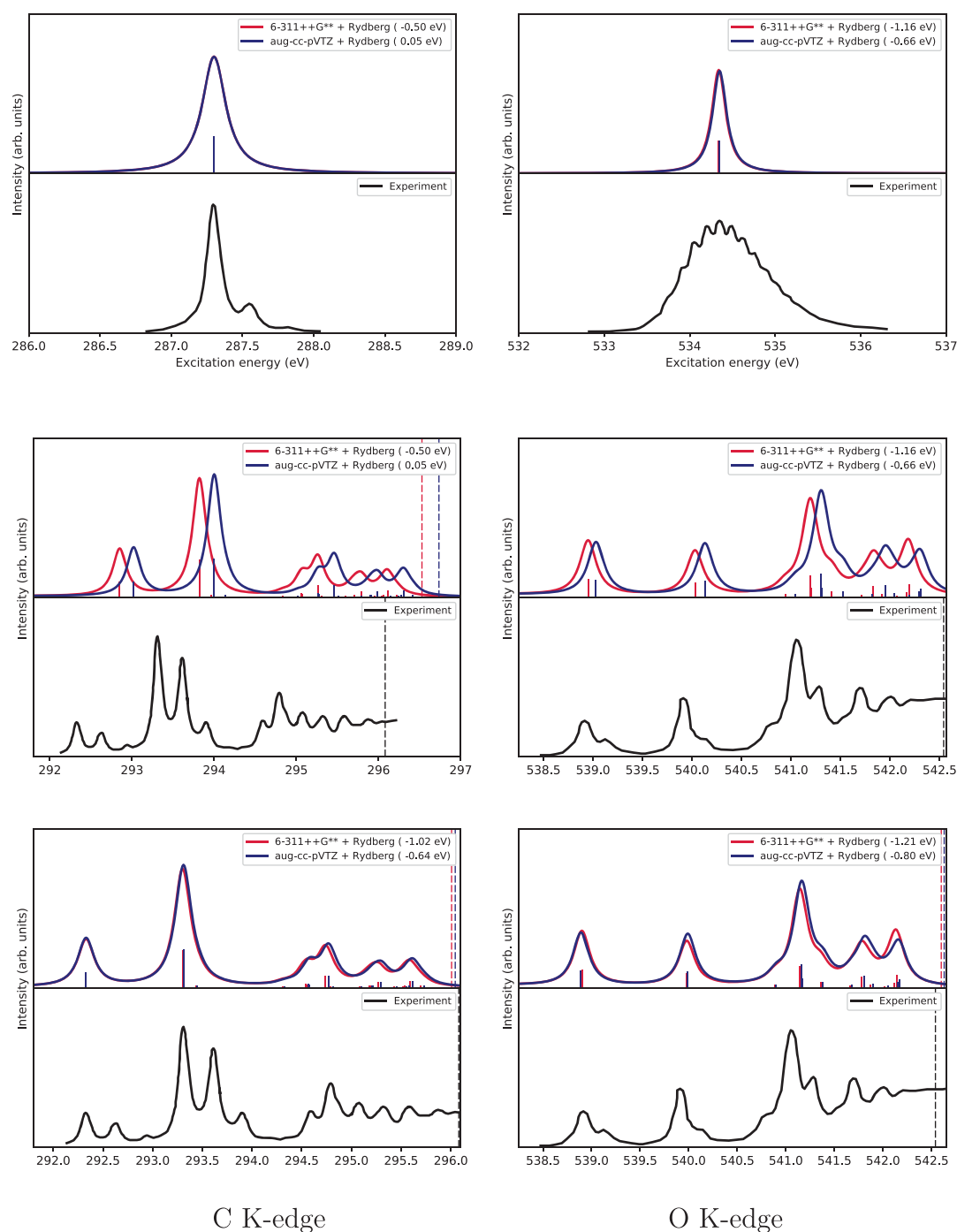


Figure 4. Carbon monoxide. fc-CVS-EOMEE-CCSD C K-edge (left) and O K-edge (right) X-ray absorption spectra obtained by convolution of the computed excitation energies and oscillator strengths with a Lorentzian function (fwhm = 0.2 eV). Upper panels show the main $1s \rightarrow \pi^*$ band; mid and bottom panels show the band progressions of the weaker $1s \rightarrow 3s\sigma$, $3p\pi$, $3d\pi$, $4s\sigma$, and $4s\pi$ transitions. Experimental spectra (0.03 eV resolution at the C K-edge and 0.07 eV resolution at the O K-edge) were digitized from ref 89. Vertical dashed lines correspond to the IEs.

Table S10 presents the spectral data for C and O K-edges of carbon monoxide, and the corresponding spectra are shown in Figure 4. The two upper panels in the figure show the main NEXAFS bands, experimentally observed between 286.5 and 289.0 eV for carbon and between 533 and 537 eV for oxygen. The middle and bottom panels of Figure 4 show the (much weaker) peaks observed at higher frequencies below the ionization limit.

The position of the dominant C K-edge $1s \rightarrow \pi^*$ band is blue-shifted by 0.50 eV in the 6-311++G** + Rydberg basis set

and red-shifted by 0.05 eV in the aug-cc-pVTZ + Rydberg basis. The O K-edge $1s \rightarrow \pi^*$ band is blue-shifted by about 1.16 eV in Pople's set and by 0.66 eV in Dunning's basis. The additional features of the main experimental bands are due to the vibronic progression, which is not included in our calculations. The overall shifts are significantly smaller than those obtained with the CVS-LR-CCSD scheme,²⁸ see also Table S2.

Upon alignment of the computed spectra with the main peak of the experimental ones, the Rydberg transitions are still

slightly misaligned; see middle panels of Figure 4. Nonetheless, all weaker $3s\sigma$, $3p\pi$, $3p\sigma$, $3d\pi$, $4s\sigma$, and $4s\pi$ transitions can be identified in the computed spectra of each edge, although, once again, without their finer vibronic progressions. The assignments can be verified by realignment of the first peak of the first progression, as shown in the bottom panels of Figure 4.

Figure 5 reports the computed spectra of ethylene obtained by convolution of the spectral data in Table S11. In this case,

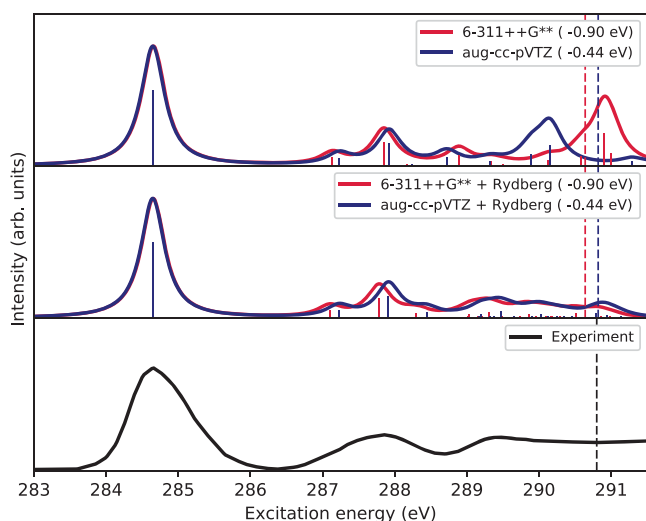


Figure 5. Ethylene. *fc*-CVS-EOMEE-CCSD X-ray absorption spectra by Lorentzian broadening (fwhm = 0.4 eV) of the computed excitation energies and oscillator strengths. Experimental spectrum (0.6 eV resolution) was digitized from ref 91. Vertical dashed lines correspond to the IEs. Energy shifts required to align the NEXAFS profiles in each basis set with the experimental one are indicated in parentheses. Computed IEs have been shifted by the same amount as used to align the NEXAFS profiles.

the Rydberg functions also improve the description of the higher energy region approaching the ionization limit (third experimental band⁹¹). The second band in the experimental spectrum corresponds to three excitations in the computed spectra. The overall shift is 0.44 eV in the aug-cc-pVTZ-(+Rydberg) set and 0.90 eV for Pople's 6-311++G**-(+Rydberg) set. The overall shifts are smaller than obtained with the CVS-LR-CCSD scheme,²⁸ see also Table S2. Upon realignment with respect to the $1s \rightarrow \pi^*$ absorption energy, the IE obtained with Pople's set is slightly underestimated compared to the experimental IEs.

Figure 6 shows the computed X-ray spectra at the C K-edge in vinylfluoride (CH_2CHF); the raw data are given in Table S12. The computed spectra were shifted to align them to the first experimental peak,⁹¹ whose position we estimated to be at 285 eV. The applied shift is -0.44 eV for Dunning's set and -0.91 eV for Pople's set. Inclusion of Rydberg-type functions in the basis set has a more modest effect than in the case of ethylene.

NTOs of the most intense core excitations obtained with the 6-311++G** basis set are shown in Table 1, allowing us to identify from which of the two C atoms they originate from and the character of the transition. The valence/Rydberg character of these intense core excitations is quantified in Table 2.

The X-ray absorption spectra obtained at the fluorine K-edge of CH_2CHF are shown in Figure 7; the raw data are given

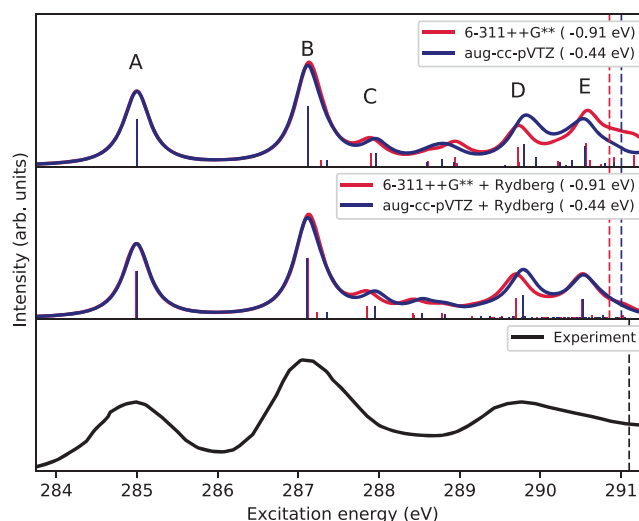


Figure 6. Vinyl fluoride. *fc*-CVS-EOMEE-CCSD X-ray absorption spectra at the C K-edge obtained by convolution of the computed energies and oscillator strengths with a Lorentzian function (fwhm = 0.4 eV). Experimental spectrum (0.6 eV resolution) was digitized from ref 91. Dashed vertical lines correspond to the IEs of the 1s electron on the carbon atom of the CH_2 group. IEs of the 1s electron of the C_{CHF} atom are outside of the displayed frequency range (experimental IE 293.48 eV). Energy shifts required to align the NEXAFS profiles in each basis set with the experimental one are indicated in parentheses. Computed IEs have been shifted by the same amount as used to align the NEXAFS profiles.

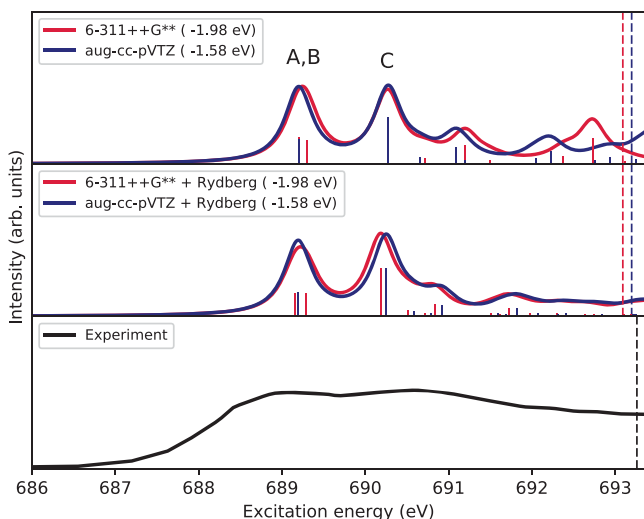
Table 1. Vinylfluoride: *fc*-CVS-EOM-CCSD/6-311++G** NTOs of 5 Selected Core-Excited States at the C K-Edge (NTO isosurface is 0.05)

State	Hole	σ_K^2	Particle
A		0.78	
B		0.79	
C		0.82	
D		0.62	
		0.19	
E		0.82	

in Table S13. In the experimental spectrum, digitized from ref 91, only two peaks are clearly discernible, with absolute energies assigned at 689.2 ± 2.0 and 690.6 ± 2.0 eV ($1s_F \rightarrow \sigma^*$ (C-F)). In the experimental study, the first peak is assigned to

Table 2. Vinylfluoride: Changes in the Second Moments of Charge Density (in Å²) and Electron and Hole Size Components (in Å) for Selected Core-Excited States at the C K-Edge

state	$\Delta\langle x^2 \rangle$	$\Delta\langle y^2 \rangle$	$\Delta\langle z^2 \rangle$	$\Delta\langle r^2 \rangle$	x_e	y_e	z_e	r_e	x_h	y_h	z_h	r_h
A	−0.35	−0.05	0.36	−0.05	1.92	0.68	0.81	1.40	0.10	0.10	0.10	0.17
B	0.36	−0.17	0.38	0.57	0.92	0.67	0.81	1.39	0.10	0.10	0.10	0.17
C	1.57	5.80	1.11	8.49	1.42	2.39	1.22	3.03	0.10	0.10	0.10	0.17
D	2.16	3.86	1.50	7.52	1.81	1.78	1.36	2.88	0.51	0.29	0.10	0.59
E	1.24	2.43	0.73	4.40	1.50	1.58	1.03	2.40	0.10	0.10	0.10	0.17

**Figure 7.** Vinylfluoride. fc-CVS-EOMEE-CCSD X-ray absorption spectra at the fluorine K edge obtained by convolution of the computed energies and oscillator strengths with a Lorentzian function (fwhm = 0.4 eV). Experimental spectrum was digitized from ref 91. Dashed vertical lines correspond to the IEs. Energy shifts required to align the NEXAFS profiles in each basis set with the experimental one are indicated in parentheses. Computed IEs have been shifted by the same amount as used to align the NEXAFS profiles. Shift was computed based on the experimentally derived maximum at 689.2 eV.

a $1s_F \rightarrow \pi^*(C=C)$ transition and the second one to a $1s_F \rightarrow \sigma^*(C-F)$ transition. Inspection of the results in Table S13 and of the NTOs in Table 3 indicates that the first band results from two almost degenerate transitions, $1s_F \rightarrow \sigma^*(C-F)$ and $1s_F \rightarrow$

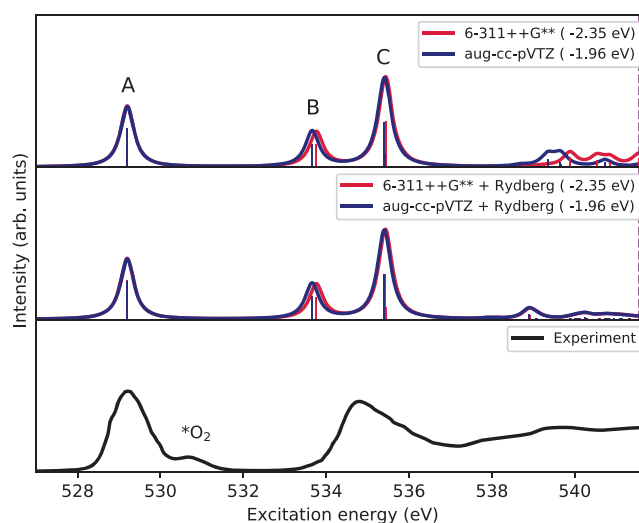
Table 3. Vinylfluoride: fc-CVS-EOMEE-CCSD/6-311++G** NTOs of 3 Selected Core-Excited States at the F K-Edge (NTO isosurface is 0.05)

State	Hole	σ_K^2	Particle
A		0.84	
B		0.79	
C		0.83	

$\pi^*(C=C)$. The third excitation (second experimental band) also appears to be of $1s_F \rightarrow \sigma^*(C-F)$ character.

The experimental IE is at 693.26 eV.⁹⁷ The computed spectra in the Pople set (with and without Rydberg functions) are shifted by −1.98 eV and those for the Dunning basis by −1.58 eV.

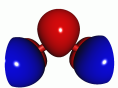
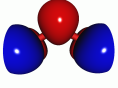
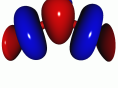
Figure 8 shows the fc-CVS-EOMEE-CCSD NEXAFS spectra of O₃ based on the spectral data in Table S14. This

**Figure 8.** Ozone. fc-CVS-EOMEE-CCSD X-ray absorption spectra obtained from convolution with a Lorentzian function (fwhm = 0.4 eV) of the computed excitation energies and oscillator strengths. Experimental spectrum high-resolution inner-shell (0.05 eV resolution) was digitized from ref 92. Dashed vertical lines correspond to the ionization energy of the terminal O atom. Central oxygen's IE was omitted as it lies above 545 eV.

molecule displays the largest overall shift relative to the experimental spectrum,⁹² −2.35 eV in the Pople set and −1.96 eV with Dunning's set. Apart from this our calculations confirm the assignment in ref 92: the first spectral feature is due to the terminal oxygens' $1s \rightarrow \pi^*$, whereas the second (broad) band is due to both the central oxygens $1s \rightarrow \pi^*$ and the terminal oxygens' $1s \rightarrow \sigma^*$ excitations; see also the NTOs in Table 4. The shoulder at 530.7 eV in the experimental spectrum is known to be due to the $1s_O \rightarrow \sigma^*$ transition of a small amount of O₂ present in the sample.⁹²

The final system considered here is adenine, whose NEXAFS and XPS spectra were experimentally recorded in the gas phase by Plekan et al.⁸⁶ We considered both carbon and nitrogen K-edges. Due to the relatively large size of this system, we only carried out calculations in the 6-311++G** basis set. The upper and lower panels of Figure 9 show the C and N K-edge spectra, respectively. The raw data are given in Table S15. The C K-edge spectra were shifted by

Table 4. Ozone: fc-CVS-EOMEE-CCSD/6-311++G NTOs of the First 3 Core-Excited States (NTO isosurface is 0.05)**

State	Hole	σ_K^2	Particle
A		0.70	
			
B		0.72	
			
C		0.63	
			

−1.10(nonplanar)/−1.14(planar) eV and the N K-edge one by −1.43(nonplanar)/−1.45(planar) eV, and one can expect an even smaller shift had the larger aug-cc-pVTZ basis set been used. The experimental features are, once again, quite well reproduced. The agreement for the higher energy peaks could probably be further improved by inclusion of Rydberg functions. Remarkably, the C K-edge spectrum obtained from the planar geometry is more similar to the experimental spectrum, primarily due to the larger splitting between the fourth and the fifth excitations in the nonplanar structure. The spectral assignment for both structures is, nonetheless, identical. This is best appreciated looking at the NTOs for the first 5 excitations shown in Tables 5 and 6. We also note that in this case, as in other examples, NTOs reveal that the electronic transitions have rather simple character and can be described by a single NTO pair. In contrast, the EOM wave functions often show multiple amplitudes with comparable weights, giving a misleading impression of the character of the transition.

4.2. Core-Level Transient Absorption Spectroscopy.

The advances in X-ray Free-Electron Lasers in the past decade have boosted the interest in computational methodologies to simulate time-resolved X-ray absorption (TR-XAS or TR-NEXAFS).^{8,9,98,99} Typically, in TR-NEXAFS pump–probe experiments the sample is first brought to a valence excited state by a UV pulse and then probed, at different time delays, with X-ray radiation. To simulate these processes, methods to compute the intensity of valence-to-core transitions are needed. An EOM-CCSD/CC3 methodology, based on the CVS approach of ref 28, has been devised and used, for instance, to simulate and interpret TR-NEXAFS experiments in thymine.⁸ The study aimed at assessing the ability of K-edge resonant absorption spectroscopy to probe ultrafast $\pi\pi^*/n\pi^*$ internal conversion in organic chromophores. Other methodologies have also been devised within the ADC,^{100–102} (MOM)-TDDFT,⁹⁹ and TP-DFT¹³ frameworks.

We extended the fc-CVS-EOMEE-CCSD formalism to the computation of the transition density matrices between two

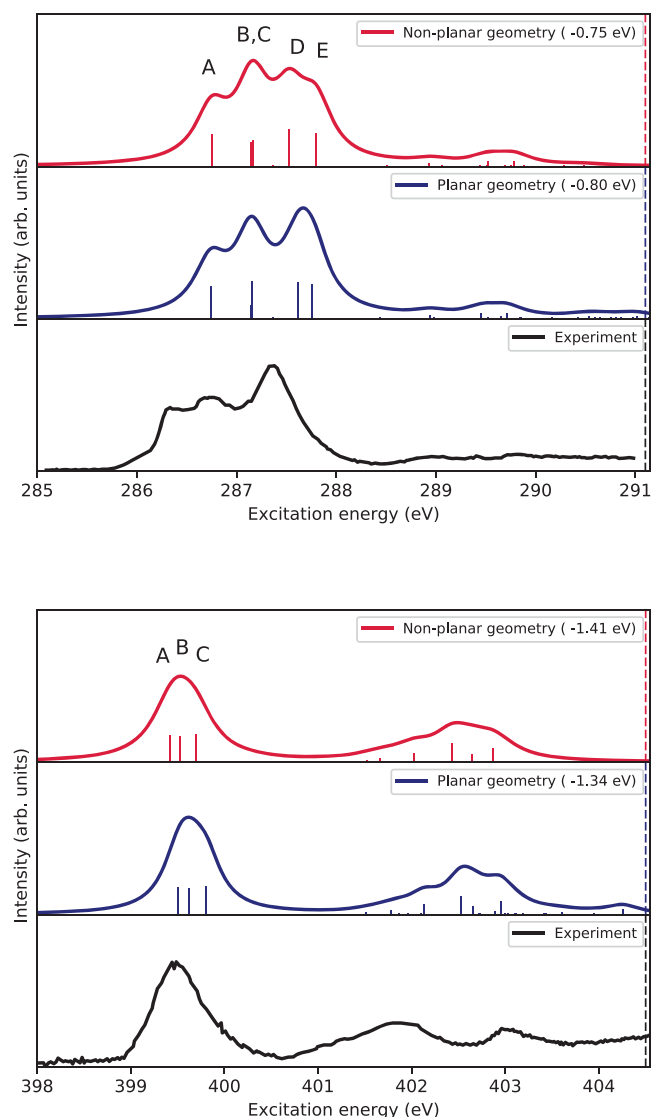


Figure 9. Adenine. C K-edge (upper panels) and N K-edge (lower panels) fc-CVS-EOMEE-CCSD/6-311++G** X-ray absorption spectra for two different molecular structures obtained by convolution of the computed energies and oscillator strengths with a Lorentzian function (fwhm = 0.4 eV). Rigid shifts applied are indicated in parentheses in the legends. They were determined with respect to the first experimental peak position in each spectrum, estimated to be at 286.4 eV for C and 399.4 eV for N. Vertical dashed line corresponds to the first IE. Computed IEs have been shifted by the same amount as used to align the NEXAFS profiles. Experimental spectra (0.57 eV resolution at the C K-edge and 0.59 eV resolution at the N K-edge) were digitized from ref 86.

excited states from which the transient X-ray absorption spectra can then be obtained. As an illustrative example, we considered the valence-to-core spectra of uracil at the O, C, and N K-edges. TR-NEXAFS spectra of uracil have not been experimentally measured yet, but they are expected to bear strong similarities with those of thymine, whose O K-edge TR-NEXAFS was measured in ref 8. Two valence excited states were considered, the first bright $\pi\pi^*$ state (S_2 at FC geometry) and the first dark $n\pi^*$ (S_1 at FC geometry) state. The NTOs of these two states, obtained at the Franck–Condon geometry, are shown in Table 7.

Table 5. Adenine: fc-CVS-EOMEE-CCSD/6-311++G** NTOs of the First 5 Core-Excited States at the C K-Edge at the Nonplanar RI-MP2/cc-pVTZ Geometry (left) and Planar B3LYP/cc-pVTZ Geometry (right) (NTO isosurface is 0.05)

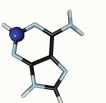
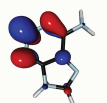
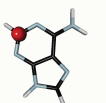
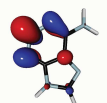
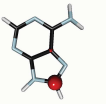
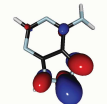
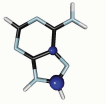
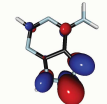
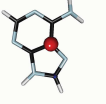
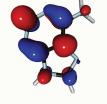
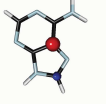
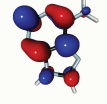
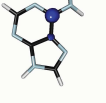
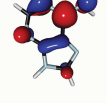
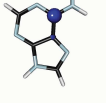
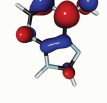
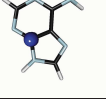
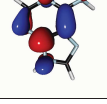
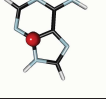
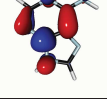
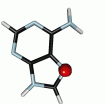
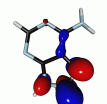
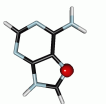
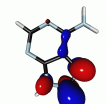
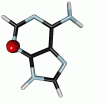
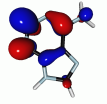
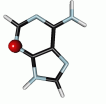
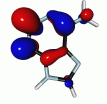
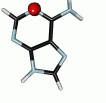
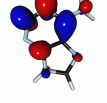
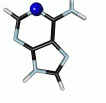
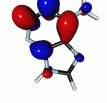
State	Hole	Non-planar σ_K^2	Particle	Hole	Planar σ_K^2	Particle
A		0.81			0.81	
B		0.79			0.71	
C		0.77			0.71	
D		0.75			0.82	
E		0.81			0.81	

Table 6. Adenine: fc-CVS-EOMEE-CCSD/6-311++G** NTOs of the First 3 Core-Excited States at the N K-Edge at the Nonplanar RI-MP2/cc-pVTZ Geometry (left) and Planar B3LYP/cc-pVTZ Geometry (NTO isosurface is 0.05)

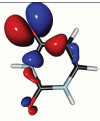
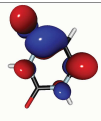
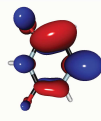
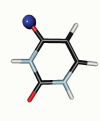
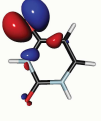
State	Hole	Non-planar σ_K^2	Particle	Hole	Planar σ_K^2	Particle
A		0.78			0.78	
B		0.78			0.78	
C		0.79			0.79	

Given the localized nature of the $n_{\text{O}}\pi^*$ (S_1) state on one of the two oxygen nuclei and similar to what has been observed for thymine,⁸ one can expect that the TR-NEXAFS measurements at the O K-edge are the best to probe the population of the $n_{\text{O}}\pi^*$ due to ultrafast internal conversion. Indeed, we show in Figures 10 and 11 the X-ray absorption spectra obtained at the O K-edge for both the ground and the two excited states at different optimized geometries for the ground and the two valence excited states. In all cases, core excitation from the $n_{\text{O}}\pi^*$ state results in the emergence of a relatively strong and distinctive signal at around 526.0–526.5 eV, similar to what has been observed for thymine.⁸ The NTO of this excitation,

labeled $1s_{\text{O}}n_{\text{O}}$, is also shown in Table 7, clearly illustrating that the core electron fills the vacancy in the $n_{\text{O}}\pi^*$ excited state.

To conclude this section, we also considered the transient state spectra that one could expect to observe if probing at the C and N K-edges after the initial pump, along with the computed ground-state NEXAFS spectra and their experimental counterparts. Figure 12 shows that at the C K-edge the valence-to-core spectra are rather weak and that, opposite to the O K-edge case, the most intense features at this edge originate from the $\pi\pi^*$ excited state. At the N K-edge (see Figure 13) the intensities of the transient absorption spectra are higher than at the C K-edge and, as in the C K-edge case, the dominant features are from the $\pi\pi^*$ excited state.

Table 7. Uracil: EOMEE-CCSD/6-311++G** NTOs of the First 2 Valence Excited States and fc-CVS-EOMEE-CCSD/6-311++G** NTO of the Core Excitation from the S_1 Valence Excited State (NTO isosurface is 0.05)

Excitation	Hole	σ_K^2	Particle
$n_O\pi^*$		0.81	
$\pi\pi^*$		0.75	
$1s_O n_O$		0.45	

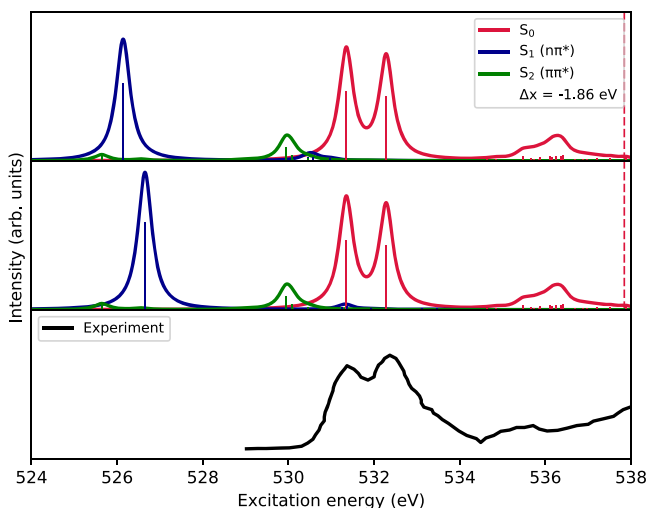


Figure 10. O K-edge of uracil. (Upper panel) fc-CVS-EOMEE-CCSD/6-311++G** ground- and excited-state core absorption spectra at the DFT Franck–Condon geometry of ref 83. (Lower panel) fc-CVS-EOMEE-CCSD/6-311++G** ground- and excited-state core-absorption spectra at the Franck–Condon geometry for both the ground state (S_0) and the $\pi\pi^*$ (S_2) states and at the TD-DFT-optimized S_1 geometry of ref 83 for S_1 . In both cases, a Lorentzian convolution function (fwhm = 0.4 eV) was used. The excited-state spectra have not been scaled to account for the population of the initial excited state.

5. CONCLUSIONS

We have presented a new, fully analytic core–valence separated equation-of-motion approach, named fc-CVS-EOM-CCSD, for calculating spectral descriptors of X-ray absorption spectroscopies, specifically near-edge absorption fine structure, core-ionization energies, and transient-state (time-resolved) X-ray absorption. The approach exploits the large energy separation of the core and valence orbitals both in the determination of the amplitudes of both coupled-cluster ground-state wave function (via the frozen-core condition) and of the EOM target-state wave functions (via CVS). By decoupling core-level states from the high-lying valence states that are embedded in the ionization continuum, CVS addresses the most challenging issue in modeling core states, that is,

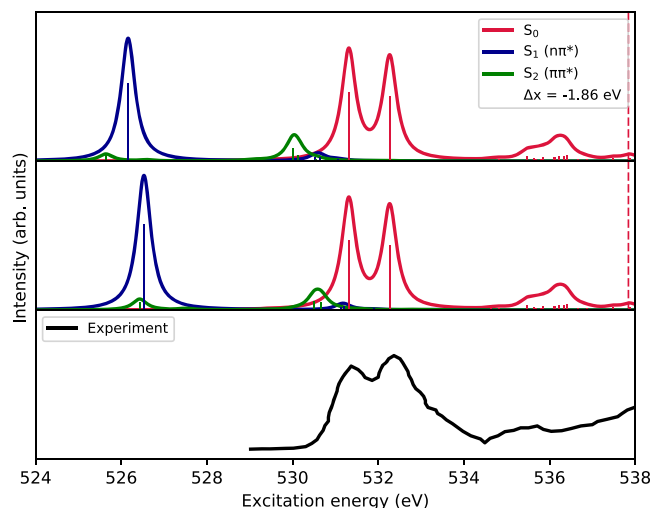


Figure 11. O K-edge of uracil. (Upper panel) fc-CVS-EOMEE-CCSD/6-311++G** ground- and excited-state core-absorption spectra at the optimized MP2/cc-pVTZ Franck–Condon geometry. (Lower panel) fc-CVS-EOMEE-CCSD/6-311++G** ground- and excited-state core-absorption spectra at planar optimized geometries for each state, i.e., MP2/cc-pVTZ for the ground state, and EOM-CCSD/aug-cc-pVDZ for the two valence excited states. In both cases, a Lorentzian convolution function (fwhm = 0.4 eV) was used. The excited-state spectra have not been scaled to account for the population of the initial excited state.

complications due to their autoionizing nature. In contrast to full EOM-CCSD, fc-CVS-EOM-CCSD features a robust numeric performance and reduced computational costs. Yet, it retains sufficient correlation, necessary to describe orbital relaxation effects, and is size intensive by construction.

We benchmarked the method using a number of atomic and molecular systems, comparing the results with both a previously proposed CVS-CCSD scheme²⁸ and experimental data. The shape of the computed NEXAFS spectra agrees very well with the experimental one in terms of the relative heights of the individual peaks and the distance between them. However, the computed spectra are shifted with respect to the experiment. The magnitude of the shifts required for the alignment varies between 0.2 and 3 eV, depending on the edge and basis set considered. In all cases, the shifts are smaller than those obtained with the previously presented CVS-CCSD approach based on the energy separation between core and valence excited states,²⁸ whereas the spectral profiles are essentially the same. Importantly, for all examples, we observed a systematic decrease of the shift upon a basis set increase.

The reduced absolute shifts from experiments compared to the previously presented CVS-CCSD approach²⁸ are most likely due to effective error cancellation between the neglect of core correlation in the ground state, its inclusion in the excited states, and the neglect of higher order correlation and excitation effects in the (EOM-)CCSD method. Even if one can argue that triple and higher excitation effects (as well as relativistic effects) are required to attain fully quantitative agreement with experiment,^{103,104} we believe that the results reported here give plenty of evidence that our ab initio approach is a useful addition to the toolbox of computational spectroscopy: it offers a robust, black-box, and reliable scheme for interpretation of modern X-ray experiments at a lower computational cost and more efficiently than the previously proposed CVS scheme,²⁸ thanks to the combination of a fully

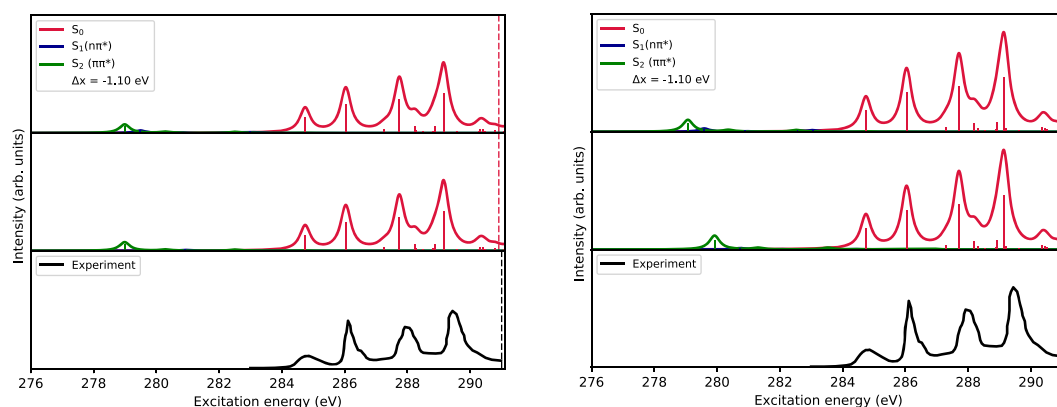


Figure 12. C K-edge of uracil. fc-CVS-EOMEE-CCSD/6-311++G** ground- and excited-state core-absorption spectra at different geometries. (Left) DFT geometries of ref 83, at the Franck–Condon DFT geometry for all states (upper panel), and the Franck–Condon DFT geometry for S_0 and S_2 and the TD-DFT-optimized S_1 geometry for S_1 (middle panel). (Right) Planar MP2/cc-pVTZ Franck–Condon geometry for all states (upper panel) and the MP2-optimized geometry for S_0 and planar-optimized EOM-CCSD/aug-cc-pVDZ for both S_1 and S_2 (middle panel). The excited-state spectra have not been scaled to account for the population of the initial excited state.

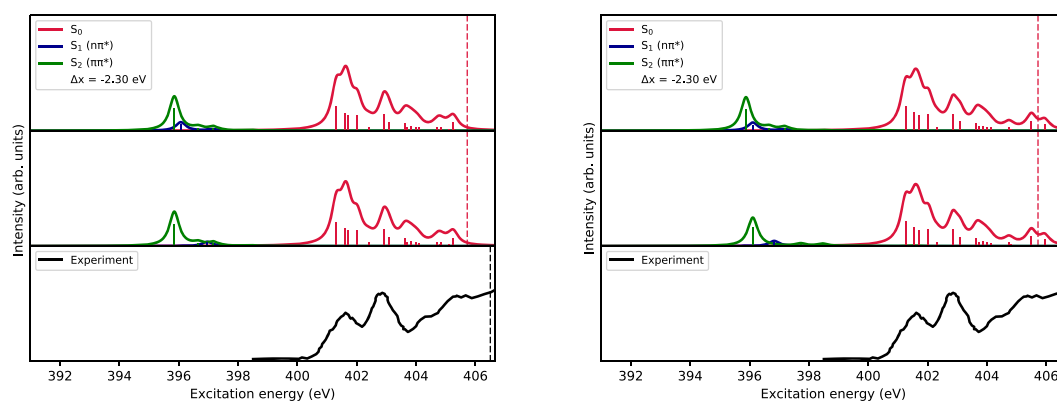


Figure 13. N K-edge of uracil. fc-CVS-EOMEE-CCSD/6-311++G** ground- and excited-state core-absorption spectra at different geometries: Franck–Condon DFT geometry of ref 83 for all states (left upper panel), Franck–Condon DFT geometry of ref 83 for S_0 and S_2 and the TD-DFT-optimized S_1 geometry for S_1 (left lower panel); planar MP2/cc-pVTZ Franck–Condon geometry for all states (right upper panel); MP2-optimized geometry for S_0 and planar optimized EOM-CCSD/aug-cc-pVDZ geometry for both S_1 and S_2 (bottom right panel). The excited-state spectra have not been scaled to account for the population of the initial excited state.

analytical implementation of the CVS condition instead of a projection technique, and the use of the frozen core approximation in the ground-state calculation.

Simulations of the transient-state NEXAFS spectra of uracil at all three edges supports the ability to probe the ultrafast internal conversion of this RNA basis by TR-NEXAFS, similar to what has been recently verified experimentally and computationally for the DNA basis thymine.⁸

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.9b00039.

Tables with raw spectral data, proof of size extensivity, comparison with CVS-LRCCSD results, and implementation formulas (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: soco@kemi.dtu.dk.

ORCID

Marta L. Vidal: 0000-0003-0653-2078

Anna I. Krylov: 0000-0001-6788-5016

Sonia Coriani: 0000-0002-4487-897X

Notes

The authors declare the following competing financial interest(s): A.I.K. is a partowner and a board member of Q-Chem, Inc.

■ ACKNOWLEDGMENTS

M.L.V. and S.C. acknowledge support from DTU Chemistry (start-up Ph.D. grant). S.C. acknowledges support from the Independent Research Fund Denmark–DFF-Forskningsprojekt2 grant no. 7014-00258B and from the H2020-MSCA-ITN-2017 training network “COSINE–Computational Spectroscopy In Natural sciences and Engineering”. A.I.K. acknowledges support by the U.S. National Science Foundation (No. CHE-1856342) and by the Simons Foundation. We thank R. Faber for useful discussion.

■ REFERENCES

(1) In *X-Ray Absorption and X-ray Emission Spectroscopy; Theory and Applications*; van Bokhoven, J., Lamberti, C., Eds.; Wiley & Sons, 2016.

- (2) In *Synchrotron Radiation: Basics, Methods and Applications*; Mobilio, S.; Boscherini, F.; Meneghini, C., Eds.; Springer, 2014.
- (3) In *X-Ray Free Electron Lasers: Applications in Materials, Chemistry and Biology*; Energy and Environment Series 18; Bergmann, U., Yachandra, V., Yano, J., Eds.; Royal Society of Chemistry, 2017.
- (4) Nisoli, M.; Decleva, P.; Calegari, F.; Palacios, A.; Martin, F. Attosecond Electron Dynamics in Molecules. *Chem. Rev.* **2017**, *117*, 10760–10825.
- (5) Norman, P.; Dreuw, A. Simulating X-ray Spectroscopies and Calculating Core-Excited States of Molecules. *Chem. Rev.* **2018**, *118*, 7208–7248.
- (6) Stöhr, J. *NEXAFS Spectroscopy*; Springer: Berlin, 1992.
- (7) Kostko, O.; Bandyopadhyay, B.; Ahmed, M. Vacuum Ultraviolet Photoionization of Complex Chemical Systems. *Annu. Rev. Phys. Chem.* **2016**, *67*, 19–40.
- (8) Wolf, T.; Myhre, R.; Cryan, J.; Coriani, S.; Squibb, R.; Battistoni, A.; Berrah, N.; Bostedt, C.; Bucksbaum, P.; Coslovich, G.; Feifel, R.; Gaffney, K.; Grilj, J.; Martinez, T.; Miyabe, S.; Moeller, S.; Mucke, M.; Natan, A.; Obaid, R.; Osipov, T.; Plekan, O.; Wang, S.; Koch, H.; Gühr, M. Probing ultrafast $\pi\pi^*/n\pi^*$ internal conversion in organic chromophores via K-edge resonant absorption. *Nature Commun.* **2017**, *8*, 1–7.
- (9) Milne, C.; Penfold, T.; Chergui, M. Recent experimental and theoretical developments in time-resolved X-ray spectroscopies. *Coord. Chem. Rev.* **2014**, *277–278*, 44–68.
- (10) Kraus, P. M.; Zürich, M.; Cushing, S. K.; Neumark, D. M.; Leone, S. R. The ultrafast X-ray spectroscopic revolution in chemical dynamics. *Nature Reviews Chemistry* **2018**, *2*, 82–94.
- (11) Besley, N.; Asmurf, F. Time-dependent density functional theory calculations of the spectroscopy of core electrons. *Phys. Chem. Chem. Phys.* **2010**, *12*, 12024–12039.
- (12) Triguero, L.; Pettersson, L. G. M.; Ågren, H. Calculations of near-edge x-ray-absorption spectra of gas-phase and chemisorbed molecules by means of density-functional and transition-potential theory. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1998**, *58*, 8097–8110.
- (13) Ehlert, C.; Gühr, M.; Saalfrank, P. An efficient first principles method for molecular pump-probe NEXAFS spectra: Application to thymine and azobenzene. *J. Chem. Phys.* **2018**, *149*, 144112.
- (14) Evangelista, F. A.; Shushkov, P.; Tully, J. C. Orthogonality Constrained Density Functional Theory for Electronic Excited States. *J. Phys. Chem. A* **2013**, *117*, 7378–7392.
- (15) Derricotte, W. D.; Evangelista, F. A. Simulation of X-ray absorption spectra with orthogonality constrained density functional theory. *Phys. Chem. Chem. Phys.* **2015**, *17*, 14360–14374.
- (16) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular electronic structure theory*; Wiley & Sons, 2000.
- (17) Koch, H.; Jørgensen, P. Coupled Cluster Response Functions. *J. Chem. Phys.* **1990**, *93*, 3333–3344.
- (18) Stanton, J.; Bartlett, R. The equation of motion coupled-cluster method. A systematic biorthogonal approach to molecular excitation energies, transition probabilities, and excited state properties. *J. Chem. Phys.* **1993**, *98*, 7029–7039.
- (19) Christiansen, O.; Jørgensen, P.; Hättig, C. Response Functions from Fourier Component Variational Perturbation Theory Applied to a Time-Averaged Quasienergy. *Int. J. Quantum Chem.* **1998**, *68*, 1.
- (20) Krylov, A. I. Equation-of-Motion Coupled-Cluster Methods for Open-Shell and Electronically Excited Species: The Hitchhiker's Guide to Fock Space. *Annu. Rev. Phys. Chem.* **2008**, *59*, 433–462.
- (21) Helgaker, T.; Coriani, S.; Jørgensen, P.; Kristensen, K.; Olsen, J.; Ruud, K. Recent Advances in Wave Function-Based Methods of Molecular-Property Calculations. *Chem. Rev.* **2012**, *112*, 543–631.
- (22) Bartlett, R. J. Coupled-cluster theory and its equation-of-motion extensions. *WIREs Comput. Mol. Sci.* **2012**, *126*–138.
- (23) Christiansen, O. Coupled Cluster Theory with Emphasis on Selected New Developments. *Theor. Chem. Acc.* **2006**, *116*, 106.
- (24) Sneskov, K.; Christiansen, O. Excited state coupled cluster methods. *WIREs Comput. Mol. Sci.* **2012**, *2*, 566–584.
- (25) Coriani, S.; Fransson, T.; Christiansen, O.; Norman, P. Asymmetric-Lanczos-Chain-Driven Implementation of Electronic Resonance Convergent Coupled-Cluster Linear Response Theory. *J. Chem. Theory Comput.* **2012**, *8*, 1616.
- (26) Peng, B.; Lestrange, P. J.; Goings, J. J.; Caricato, M.; Li, X. Energy-Specific Equation-of-Motion Coupled-Cluster Methods for High-Energy Excited States: Application to K-Edge X-Ray Absorption Spectroscopy. *J. Chem. Theory Comput.* **2015**, *11*, 4146.
- (27) Coriani, S.; Christiansen, O.; Fransson, T.; Norman, P. Coupled-Cluster Response Theory for Near-Edge X-Ray-Absorption Fine Structure of Atoms and Molecules. *Phys. Rev. A: At., Mol., Opt. Phys.* **2012**, *85*, 022507.
- (28) Coriani, S.; Koch, H. Communication: X-ray Absorption Spectra and Core-Ionization Potentials within a Core-Valence Separated Coupled Cluster Framework. *J. Chem. Phys.* **2015**, *143*, 181103.
- (29) Krylov, A. I. In *Reviews in Computational Chemistry*; Parrill, A., Lipkowitz, K., Eds.; J. Wiley & Sons, 2017; Vol. 30; Chapter 4, pp 151–224.
- (30) Sadybekov, A.; Krylov, A. I. Coupled-cluster based approach for core-ionized and core-excited states in condensed phase: Theory and application to different protonated forms of aqueous glycine. *J. Chem. Phys.* **2017**, *147*, 014107.
- (31) Hršak, D.; Nørby, M. S.; Coriani, S.; Kongsted, J. One-Photon Absorption Properties from a Hybrid Polarizable Density Embedding/Complex Polarization Propagator Approach for Polarizable Solutions. *J. Chem. Theory Comput.* **2018**, *14*, 2145–2154.
- (32) Reinhardt, W. P. Complex coordinates in the theory of atomic and molecular structure and dynamics. *Annu. Rev. Phys. Chem.* **1982**, *33*, 223–255.
- (33) Jagau, T.-C.; Bravaya, K. B.; Krylov, A. I. Extending Quantum Chemistry of Bound States to Electronic Resonances. *Annu. Rev. Phys. Chem.* **2017**, *68*, 525–553.
- (34) Emrich, K. An extension of the coupled-cluster formalism to excited states (I). *Nucl. Phys. A* **1981**, *A351*, 379–396.
- (35) Gilbert, A.; Besley, N.; Gill, P. Self-consistent field calculations of excited states using the maximum overlap method (MOM). *J. Phys. Chem. A* **2008**, *112*, 13164–13171.
- (36) Zuev, D.; Vecharynski, E.; Yang, C.; Orms, N.; Krylov, A. I. New algorithms for iterative matrix-free eigensolvers in quantum chemistry. *J. Comput. Chem.* **2015**, *36*, 273–284.
- (37) Zuev, D.; Bravaya, K. B.; Crawford, T. D.; Lindh, R.; Krylov, A. I. Electronic structure of the two isomers of the anionic form of p-coumaric acid chromophore. *J. Chem. Phys.* **2011**, *134*, 034310.
- (38) Moiseyev, N. *Non-Hermitian quantum mechanics*; Cambridge University Press, 2011.
- (39) Cederbaum, L. S.; Domcke, W.; Schirmer, J. Many-body theory of core holes. *Phys. Rev. A: At., Mol., Opt. Phys.* **1980**, *22*, 206–222.
- (40) Barth, A.; Schirmer, J. Theoretical core-level excitation spectra of N₂ and CO by a new polarisation propagator method. *J. Phys. B: At. Mol. Phys.* **1985**, *18*, 867–885.
- (41) Wenzel, J.; Wormit, M.; Dreuw, A. Calculating Core-Level Excitations and X-Ray Absorption Spectra of Medium-Sized Closed-Shell Molecules with the Algebraic-Diagrammatic Construction Scheme for the Polarization Propagator. *J. Comput. Chem.* **2014**, *35*, 1900.
- (42) Wenzel, J.; Holzer, A.; Wormit, M.; Dreuw, A. Analysis and Comparison of CVS-ADC Approaches up to Third Order for the Calculation of Core-Excited States. *J. Chem. Phys.* **2015**, *142*, 214104.
- (43) Feshbach, H. A unified theory of nuclear reactions. 2. *Ann. Phys. (Amsterdam, Neth.)* **1962**, *19*, 287–313.
- (44) Fano, U. Configuration interaction on intensities and phase shifts. *Phys. Rev.* **1961**, *124*, 1866–1878.
- (45) Fransson, T.; Coriani, S.; Christiansen, O.; Norman, P. Carbon X-ray Absorption Spectra of Fluoroethenes and Acetone: A Study at the Coupled Cluster, Density Functional, and Static-Exchange levels of Theory. *J. Chem. Phys.* **2013**, *138*, 124311.

- (46) Löwdin, P.-O. Studies in perturbation theory. IV. Solution of eigenvalue problem by projection operator formalism. *J. Math. Phys.* **1962**, *3*, 969–982.
- (47) Averbukh, V.; Cederbaum, L. Ab initio calculation of interatomic decay rates by a combination of the Fano ansatz, Green's-function methods, and the Stieltjes imaging technique. *J. Chem. Phys.* **2005**, *123*, 204107.
- (48) Sekino, H.; Bartlett, R. J. A linear response, coupled-cluster theory for excitation energy. *Int. J. Quantum Chem.* **1984**, *26*, 255–265.
- (49) We use the standard convention that $i, j, k, l, \dots$ refer to occupied orbitals, $a, b, c, d, \dots$ represent virtual ones, while $p, q, r, s, \dots$ denote generic orbitals, either occupied or unoccupied.
- (50) Stanton, J.; Gauss, J. Analytic energy derivatives for ionized states described by the equation-of-motion coupled cluster method. *J. Chem. Phys.* **1994**, *101*, 8938–8944.
- (51) Stanton, J.; Gauss, J. A discussion on some problems associated with the quantum mechanical treatment of open-shell molecules. *Adv. Chem. Phys.* **2003**, *125*, 101–146.
- (52) Pieniazek, P. A.; Bradforth, S. E.; Krylov, A. I. Charge localization and Jahn-Teller distortions in the benzene dimer cation. *J. Chem. Phys.* **2008**, *129*, 074104.
- (53) Löwdin, P.-O. Some Aspects on the Hamiltonian and Liouvillian Formalism, the Special Propagator Methods, and the Equation of Motion Approach. *Adv. Quantum Chem.* **1985**, *17*, 285–334.
- (54) Levchenko, S. V.; Krylov, A. I. Equation-of-motion spin-flip coupled-cluster model with single and double substitutions: Theory and application to cyclobutadiene. *J. Chem. Phys.* **2004**, *120*, 175–185.
- (55) Davidson, E. R. The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices. *J. Comput. Phys.* **1975**, *17*, 87–94.
- (56) Hirao, K.; Nakatsuji, H. A generalization of the Davidson's method to large nonsymmetric eigenvalue problems. *J. Comput. Phys.* **1982**, *45*, 246–254.
- (57) Rettrup, S. An iterative method for calculating several of the extreme eigensolutions of large real non-symmetric matrices. *J. Comput. Phys.* **1982**, *45*, 100–107.
- (58) Stanton, J. F. 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.* **1993**, *99*, 8840–8847.
- (59) Stanton, J. F.; Gauss, J. Analytic energy gradients for the equation-of-motion coupled-cluster method: Implementation and application to the HCN/HNC system. *J. Chem. Phys.* **1994**, *100*, 4695–4698.
- (60) Levchenko, S. V.; Wang, T.; Krylov, A. I. Analytic gradients for the spin-conserving and spin-flipping equation-of-motion coupled-cluster models with single and double substitutions. *J. Chem. Phys.* **2005**, *122*, 224106.
- (61) Nanda, K.; Krylov, A. I.; Gauss, J. Communication: The pole structure of the dynamical polarizability tensor in equation-of-motion coupled-cluster theory. *J. Chem. Phys.* **2018**, *149*, 141101.
- (62) Luzanov, A.; Sukhorukov, A.; Umanskii, V. Application of transition density matrix for analysis of excited states. *Theor. Exp. Chem.* **1976**, *10*, 354–361.
- (63) Luzanov, A.; Zhikol, O. In *Practical aspects of computational chemistry I: An overview of the last two decades and current trends*; Leszczynski, J., Shukla, M., Eds.; Springer, 2012; pp 415–449.
- (64) Plasser, F.; Lischka, H. Analysis of excitonic and charge transfer interactions from quantum chemical calculations. *J. Chem. Theory Comput.* **2012**, *8*, 2777–2789.
- (65) Plasser, F.; Wormit, M.; Dreuw, A. New tools for the systematic analysis and visualization of electronic excitations. I. Formalism. *J. Chem. Phys.* **2014**, *141*, 024106–13.
- (66) Plasser, F.; Bäppler, S.; Wormit, M.; Dreuw, A. New tools for the systematic analysis and visualization of electronic excitations. II. Applications. *J. Chem. Phys.* **2014**, *141*, 024107–12.
- (67) Martin, R. Natural transition orbitals. *J. Chem. Phys.* **2003**, *118*, 4775–4777.
- (68) Bäppler, S.; Plasser, F.; Wormit, M.; Dreuw, A. Exciton analysis of many-body wave functions: Bridging the gap between the quasiparticle and molecular orbital pictures. *Phys. Rev. A: At., Mol., Opt. Phys.* **2014**, *90*, 052521.
- (69) Mewes, S.; Plasser, F.; Krylov, A. I.; Dreuw, A. Benchmarking excited-state calculations using exciton properties. *J. Chem. Theory Comput.* **2018**, *14*, 710–725.
- (70) Schirmer, J. Beyond the random-phase approximation: A new approximation scheme for the polarization propagator. *Phys. Rev. A: At., Mol., Opt. Phys.* **1982**, *26*, 2395–2416.
- (71) Dreuw, A.; Wormit, M. The algebraic diagrammatic construction scheme for the polarization propagator for the calculation of excited states. *WIREs Comput. Mol. Sci.* **2015**, *5*, 82–95.
- (72) Angonoa, G.; Walter, O.; Schirmer, J. Theoretical K-shell ionization spectra of N₂ and CO by a fourth-order Greens function method. *J. Chem. Phys.* **1987**, *87*, 6789–6801.
- (73) Schirmer, J.; Thiel, A. An intermediate state representation approach to K-shell ionization in molecules. I. Theory. *J. Chem. Phys.* **2001**, *115*, 10621.
- (74) Purvis, G.; Bartlett, R. A full coupled-cluster singles and doubles model: the inclusion of disconnected triples. *J. Chem. Phys.* **1982**, *76*, 1910–1918.
- (75) Stanton, J.; Gauss, J. Perturbative treatment of the similarity transformed Hamiltonian in equation-of-motion coupled-cluster approximations. *J. Chem. Phys.* **1995**, *103*, 1064–1076.
- (76) Wang, Z.; Tu, Z.; Wang, F. Equation-of-Motion Coupled-Cluster Theory for Excitation Energies of Closed-Shell Systems with Spin-Orbit Coupling. *J. Chem. Theory Comput.* **2014**, *10*, 5567–5576.
- (77) Krylov, A. I.; Gill, P. M. W. Q-Chem: An engine for innovation. *WIREs Comput. Mol. Sci.* **2013**, *3*, 317–326.
- (78) Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X.; Ghosh, D.; Goldey, M.; Horn, P. R.; Jacobson, L. D.; Kaliman, I.; Khaliullin, R. Z.; Kus, T.; Landau, A.; Liu, J.; Proynov, E. I.; Rhee, Y. M.; Richard, R. M.; Rohrdanz, M. A.; Steele, R. P.; Sundstrom, E. J.; Woodcock, H. L., III; Zimmerman, P. M.; Zuev, D.; Albrecht, B.; Alguire, E.; Austin, B.; Beran, G. J. O.; Bernard, Y. A.; Berquist, E.; Brandhorst, K.; Bravaya, K. B.; Brown, S. T.; Casanova, D.; Chang, C.-M.; Chen, Y.; Chien, S. H.; Closser, K. D.; Crittenden, D. L.; Didenhofen, M.; DiStasio, R. A., Jr.; Do, H.; Dutoi, A. D.; Edgar, R. G.; Fatehi, S.; Fusti-Molnar, L.; Ghysels, A.; Golubeva-Zadorozhnaya, A.; Gomes, J.; Hanson-Heine, M. W.; Harbach, P. H.; Hauser, A. W.; Hohenstein, E. G.; Holden, Z. C.; Jagau, T.-C.; Ji, H.; Kaduk, B.; Khistyayev, K.; Kim, J.; Kim, J.; King, R. A.; Klunzinger, P.; Kosenkov, D.; Kowalczyk, T.; Krauter, C. M.; Lao, K. U.; Laurent, A. D.; Lawler, K. V.; Levchenko, S. V.; Lin, C. Y.; Liu, F.; Livshits, E.; Lochan, R. C.; Luenser, A.; Manohar, P.; Manzer, S. F.; Mao, S.-P.; Mardrossian, N.; Marenich, A. V.; Maurer, S. A.; Mayhall, N. J.; Neuscamman, E.; Oana, C. M.; Olivares-Amaya, R.; O'Neill, D. P.; Parkhill, J. A.; Perrine, T. M.; Peverati, R.; Prociuk, A.; Rehn, D. R.; Rosta, E.; Russ, N. J.; Sharada, S. M.; Sharma, S.; Small, D. W.; Sodt, A.; Stein, T.; Stück, D.; Su, Y.-C.; Thom, A. J.; Tsuchimochi, T.; Vanovschi, V.; Vogt, L.; Vydrov, O.; Wang, T.; Watson, M. A.; Wenzel, J.; White, A.; Williams, C. F.; Yang, J.; Yeganeh, S.; Yost, S. R.; You, Z.-Q.; Zhang, I. Y.; Zhang, X.; Zhao, Y.; Brooks, B. R.; Chan, G. K.; Chipman, D. M.; Cramer, C. J.; Goddard, W. A., III; Gordon, M. S.; Hehre, W. J.; Klamt, A.; Schaefer, H. F., III; Schmidt, M. W.; Sherrill, C. D.; Truhlar, D. G.; Warshel, A.; Xu, X.; Aspuru-Guzik, A.; Baer, R.; Bell, A. T.; Besley, N. A.; Chai, J.-D.; Dreuw, A.; Dunietz, B. D.; Furlani, T. R.; Gwaltney, S. R.; Hsu, C.-P.; Jung, Y.; Kong, J.; Lambrecht, D. S.; Liang, W.; Ochsenfeld, C.; Rassolov, V. A.; Slipchenko, L. V.; Subotnik, J. E.; Van Voorhis, T.; Herbert, J. M.; Krylov, A. I.; Gill, P. M.; Head-Gordon, M. Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* **2015**, *113*, 184–215.
- (79) Epifanovsky, E.; Wormit, M.; Kus, T.; Landau, A.; Zuev, D.; Khistyayev, K.; Manohar, P. U.; Kaliman, I.; Dreuw, A.; Krylov, A. I. New implementation of high-level correlated methods using a general

block-tensor library for high-performance electronic structure calculations. *J. Comput. Chem.* **2013**, *34*, 2293–2309.

(80) Stanton, J. F.; Gauss, J.; Cheng, L.; Harding, M. E.; Matthews, D. A.; Szalay, P. G. *CFOUR, Coupled-Cluster techniques for Computational Chemistry, a quantum-chemical program package*; with contributions from A. A. Auer, R. J. Bartlett, U. Benedikt, C. Berger, D. E. Bernholdt, Y. J. Bomble, O. Christiansen, F. Engel, R. Faber, M. Heckert, O. Heun, M. Hilgenberg, C. Huber, T.-C. Jagau, D. Jonsson, J. Jusélius, T. Kirsch, K. Klein, W. J. Lauderdale, F. Lipparini, T. Metzroth, L. A. Mück, D. P. O'Neill, D. R. Price, E. Prochnow, C. Puzzarini, K. Ruud, F. Schiffmann, W. Schwalbach, C. Simmons, S. Stopkowicz, A. Tajti, J. Vázquez, F. Wang, J. D. Watts and the integral packages MOLECULE (J. Almlöf and P.R. Taylor), PROPS (P. R. Taylor), ABACUS (T. Helgaker, H. J. Aa. Jensen, P. Jørgensen, and J. Olsen), and ECP routines by A. V. Mitin and C. van Wüllen. For the current version, see <http://www.cfour.de>.

(81) Khani, S. K.; Faber, R.; Santoro, F.; Hättig, C.; Coriani, S. UV Absorption and Magnetic Circular Dichroism Spectra of Purine, Adenine, and Guanine: a Coupled Cluster Study in Vacuo and in Aqueous Solution. *J. Chem. Theory Comput.* **2019**, *15*, 1242–1254.

(82) Santoro, F.; Improtà, R.; Fahleson, T.; Kauczor, J.; Norman, P.; Coriani, S. Relative Stability of the L_a and L_b Excited States in Adenine and Guanine: Direct Evidence from TD-DFT Calculations of MCD Spectra. *J. Phys. Chem. Lett.* **2014**, *5*, 1806–1811.

(83) Zhang, X.; Herbert, J. M. Excited-State Deactivation Pathways in Uracil versus Hydrated Uracil: Solvatochromatic Shift in the $1n\pi^*$ State is the Key. *J. Phys. Chem. B* **2014**, *118*, 7806–7817.

(84) Kaufmann, K.; Baumeister, W.; Jungen, M. Universal Gaussian basis sets for an optimum representation of Rydberg and continuum wavefunctions. *J. Phys. B: At., Mol. Opt. Phys.* **1989**, *22*, 2223.

(85) Reisler, H.; Krylov, A. I. Interacting Rydberg and valence states in radicals and molecules: Experimental and theoretical studies. *Int. Rev. Phys. Chem.* **2009**, *28*, 267–308.

(86) Plekan, O.; Feyer, V.; Richter, R.; Coreno, M.; de Simone, M.; Prince, K.; Trofimov, A.; Gromov, E.; Zaytseva, I.; Schirmer, J. A theoretical and experimental study of the near edge X-ray absorption fine structure (NEXAFS) and X-ray photoelectron spectra (XPS) of nucleobases: Thymine and adenine. *Chem. Phys.* **2008**, *347*, 360–375.

(87) Coreno, M.; Avaldi, L.; Camilloni, R.; Prince, K. C.; de Simone, M.; Karvonen, J.; Colle, R.; Simonucci, S. Measurement and ab initio calculation of the Ne photoabsorption spectrum in the region of the K edge. *Phys. Rev. A: At., Mol., Opt. Phys.* **1999**, *59*, 2494.

(88) Schirmer, J.; Trofimov, A.; Randall, K.; Feldhaus, J.; Bradshaw, A. M.; Ma, Y.; Chen, C. T.; Sette, F. K-shell excitation of the water, ammonia and methane molecules using high-resolution photo-absorption spectroscopy. *Phys. Rev. A: At., Mol., Opt. Phys.* **1993**, *47*, 1136.

(89) Ma, Y.; Chen, C.; Meigs, G.; Randall, K.; Sette, F. High-resolution K-shell photoabsorption measurements of simple molecules. *Phys. Rev. A: At., Mol., Opt. Phys.* **1991**, *44*, 1848.

(90) Püttner, R.; Dominguez, I.; Morgan, T.; Cisneros, C.; Fink, R.; Rotenberg, E.; Warwick, T.; Domke, M.; Kaindl, G.; Schlachter, A. Vibrationally resolved O 1s core-excitation spectra of CO and NO. *Phys. Rev. A: At., Mol., Opt. Phys.* **1999**, *59*, 3415.

(91) McLaren, R.; Clark, S.; Ishii, I.; Hitchcock, A. Absolute oscillator strengths from K-shell electron-energy-loss spectra of the fluoroethenes and 1, 3-perfluorobutadiene. *Phys. Rev. A: At., Mol., Opt. Phys.* **1987**, *36*, 1683.

(92) Stranges, S.; Alagia, M.; Fronzoni, G.; Decleva, P. High-Resolution Inner-Shell Photoabsorption and Dissociation of Ozone. *J. Phys. Chem. A* **2001**, *105*, 3400–3406.

(93) Feyer, V.; Plekan, O.; Richter, R.; Coreno, M.; de Simone, M.; Prince, K. C.; Trofimov, A. B.; Zaytseva, I. L.; Schirmer, J. Tautomerism in Cytosine and Uracil: A Theoretical and Experimental X-ray Absorption and Resonant Auger Study. *J. Phys. Chem. A* **2010**, *114*, 10270–10276.

(94) Rohatgi, A. *WebPlotDigitizer*, version 4.0; 2017; <https://automeris.io/WebPlotDigitizer>.

(95) Schaftenaar, G.; Noordik, J. Molden: a pre- and post-processing program for molecular and electronic structures. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 123–134.

(96) 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.; Hettner, H.; Hjertenæs, E.; Høst, S.; Hoyvik, I.-M.; Iozzi, M. F.; Jansik, 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.; Pawłowski, F.; Pedersen, T. B.; Provasi, P. F.; Reine, S.; Rinkevicius, Z.; Ruden, T. A.; Ruud, K.; Rybak, V. V.; Salek, 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 Comput. Mol. Sci.* **2014**, *4*, 269.

(97) Jolly, W.; Bomben, K.; Eyermann, C. Core-electron binding energies for gaseous atoms and molecules. *At. Data Nucl. Data Tables* **1984**, *31*, 433–493.

(98) Li, Z.; El-Amine Madjet, M.; Vendrell, O.; Santra, R. Core-level transient absorption spectroscopy as a probe of electron hole relaxation in photoionized $H^+(H_2O)_n$. *Faraday Discuss.* **2014**, *171*, 457–470.

(99) Bhattacharjee, A.; Pemmaraju, C. D.; Schnorr, K.; Attar, A. R.; Leone, S. R. Ultrafast Intersystem Crossing in Acetylacetone via Femtosecond X-ray Transient Absorption at the Carbon K-Edge. *J. Am. Chem. Soc.* **2017**, *139*, 16576–16583.

(100) Neville, S. P.; Averbukh, V.; Patchkovskii, S.; Ruberti, M.; Yun, R.; Chergui, M.; Stolow, A.; Schuurman, M. S. Beyond structure: ultrafast X-ray absorption spectroscopy as a probe of non-adiabatic wavepacket dynamics. *Faraday Discuss.* **2016**, *194*, 117–145.

(101) Neville, S. P.; Averbukh, V.; Ruberti, M.; Yun, R.; Patchkovskii, S.; Chergui, M.; Stolow, A.; Schuurman, M. S. Excited state X-ray absorption spectroscopy: Probing both electronic and structural dynamics. *J. Chem. Phys.* **2016**, *145*, 144307.

(102) Neville, S. P.; Chergui, M.; Stolow, A.; Schuurman, M. S. Ultrafast X-Ray Spectroscopy of Conical Intersections. *Phys. Rev. Lett.* **2018**, *120*, 243001.

(103) Myhre, R.; Wolf, T.; Cheng, L.; Nandi, S.; Coriani, S.; Gühr, M.; Koch, H. A theoretical and experimental benchmark study of core-excited states in nitrogen. *J. Chem. Phys.* **2018**, *148*, 064106.

(104) Liu, J.; Matthews, D.; Coriani, S.; Cheng, L. Benchmark Calculations of K-Edge Ionization Energies for First-Row Elements Using Scalar-Relativistic Core-Valence-Separated Equation-of-Motion Coupled-Cluster Methods. *J. Chem. Theory Comput.* **2019**, *15*, 1642–1651.