Accurate core-excited states via inclusion of core triple excitations in similarity-transformed EOM theory

Megan Simons* and Devin A. Matthews

Department of Chemistry, Southern Methodist University, Dallas, TX 75275

E-mail: msimons@smu.edu

Abstract

The phenomenon of orbital relaxation upon excitation of core electrons is a major problem in the linear-response treatment of core-hole spectroscopies. Rather than addressing relaxation through direct dynamical correlation of the excited state via equation-of-motion coupled cluster theory (EOMEE-CC), we extend the alternative similarity-transformed equation-of-motion coupled cluster theory (STEOMEE-CC) by including the core-valence separation (CVS) and correlation of triple excitations only within the calculation of core ionization energies. This new method, CVS-STEOMEE-CCSD+cT, significantly improves on CVS-EOMEE-CCSD and unmodified CVS-STEOMEE-CCSD when compared to full CVS-EOM-CCSDT for K-edge core-excitation energies of a set of small molecules. The improvement in both absolute and relative (shifted) peak positions is nearly as good as for transition-potential coupled cluster (TP-CC), which includes an explicit treatment of orbital relaxation, and CVS-EOMEE-CCSD*, which includes a perturbative treatment of triple excitations.

1 Introduction

Coupled cluster (CC) theory is one of the most powerful methods for treating dynamical correlation in molecules, and is capable of computing highly accurate ground state energies of small molecules.[1] Beyond the ground state, coupled cluster has been extended to excited states via the time-dependent linear-response formalism,[2,3] the closely related equation-of-motion approach,[4–6] as well as other approaches such as the symmetry-adapted cluster technique.[7,8] Equation-of-motion coupled-cluster (EOM-CC) theory can be used for excited (EE-EOM-CC), electron-attached (EA-EOM-CC), and ionized state (IP-EOM-CC) energies, as well as multiply-ionized/attached states and even spin-flip excitations (SF-EOM-CC).[9,10] EOM-CC is able to do this by taking advantage of the similarity transformation of the Hamiltonian, which guarantees size-extensivity of the excited state total energy (although excitation energies are not in general size-consistent with respect to charge separation). While the ground state CC wavefunction is single reference, the CI-like nature of the EOM-CC wavefunction can capture significant multi-reference character of excited states.[11,12]

Nooijen et al. proposed an alternative approach to EOM-CC of using a second similarity transformation of the Hamiltonian, followed by diagonalization in a small (CIS-like) excitation space.[13,14] In this similarity-transformed equation of motion coupled cluster (STEOM-CC) theory, the second similarity transformation simultaneously captures the dynamical correlation of all low-lying excited states. In comparison, EOM-CC determines the wavefunction for a single excited state and incorporates dynamical correlation effects via explicit inclusion of higher excitations in the excited state wavefunction. The STEOM-CC similarity transformation uses the wavefunctions of a number of ionized and electron-attached states in order to build the transformation. This transformation can be viewed as decoupling the "active" single-electron excitations from the double excitations, much as the ground state coupled cluster equations decouple the reference from single and double excitations. Thus, while STEOM-CC obtains the excited state energies by diagonalization only in the space of single excitations, it achieves an accuracy much greater than that of CIS or EOM-CCS.

Both EOMEE-CCSD and STEOMEE-CCSD have been highly successful at describing valence excited states of predominately single-excitation character, but some modifications to the theory are necessary for application in the x-ray regime. X-ray excited states, necessary for simulating spectra such as NEXAFS, XES, and RIXS,[15] are not bound states, but are resonances embedded deep in the valence continuum. For EOM-CC, the core-valence separation (CVS) approach of Coriani and Koch[16] has been highly successful, although it is also possible to employ other techniques such as damped response,[17] complex scaling/complex absorption potential,[18–20] etc. However, the CVS does not address the other major issue encountered in the x-ray regime: orbital relaxation. Due to the presence of a core hole in the excited state wavefunction, the valence orbitals undergo considerable contraction and rotation. CVS-EOMEE-CCSD incompletely captures this effect and hence overestimates core vertical excitation energies by 1–3 eV (*vide infra*). The addition of triple excitations, either full CVS-EOMEE-CCSDT[21] or an approximate treatment of triples,[22] is necessary to fully treat the relaxation effects within standard EOM-CC theory.

The CVS can be ported from EOM-CC to STEOM-CC in a relatively straightforward manner.[23] However, STEOM-CC also offers an alternative approach to the orbital relaxation issue. Here we propose a modification of CVS-STEOMEE-CCSD which efficiently and accurate treats orbital relaxation for core excited states, which we dub CVS-STEOMEE-CCSD+cT. This method compares favorably with standard CVS-(ST)EOM-CCSD, and achieves similar performance compared to an approximate inclusion of triples (CVS-EOMEE-CCSD*), as well as to an explicit inclusion of relaxation effects in the reference via TP-CCSD.[24]

2 Theoretical Methods

2.1 EOM-CC

The coupled cluster ground state[25] is characterized by a non-hermitian similarity transformation of the Hamiltonian which decouples the reference from the space of excited determinants,

$$\bar{H} = e^{-\hat{T}} H e^{\hat{T}}, \tag{1}$$

$$\hat{T} = \sum_{k=1}^{N} \hat{T}_k = \sum_{ai} t_i^a a_a^\dagger a_i + \frac{1}{4} \sum_{abij} t_{ij}^{ab} a_a^\dagger a_b^\dagger a_j a_i + \cdots, \tag{2}$$

$$\langle \Phi_0 | \bar{H} | \Phi_0 \rangle = E_{CC}, \tag{3}$$

$$\langle \Phi_{i_1 \ldots i_k}^{a_1 \ldots a_k} | \bar{H} | \Phi_0 \rangle = 0, \quad 0 < k \leq N, \tag{4}$$

where $|\Phi_0\rangle$ is a zeroth order description of the wave function (single determinant) and $|\Phi_{i_1 \ldots i_k}^{a_1 \ldots a_k}\rangle$ are the kth excited determinants. In EOM-CCSD, the CC equations are solved in the space of single and double excitations and hence $N = 2$.

The transformed Hamiltonian $\bar{H}$, defined via the cluster operators $\hat{T}_k$, provides the means to obtain excited states as well. The ground state coupled cluster energy is an eigenvalue of $\bar{H}$ with distinct right and left eigenfunctions due to the non-hermitian nature of the similarity transformation,

$$\bar{H} \hat{R}(0) | \Phi_0 \rangle = E_{CC} \hat{R}(0) | \Phi_0 \rangle, \quad \hat{R}(0) = \hat{1}, \tag{5}$$

$$\langle \Phi_0 | \hat{L}(0) \bar{H} = \langle \Phi_0 | \hat{L}(0) E_{CC}, \quad \hat{L}(0) = \hat{1} + \hat{\Lambda}. \tag{6}$$

Explicit diagonalization of the transformed Hamiltonian, shifted by the ground state energy, yields the vertical excitation energies ω_i and their associated right and left eigen-

functions,[6]

$$\left(\bar{H} - E_{CC}\right) \hat{R}(m)|\Phi_0\rangle = [\bar{H}, \hat{R}(m)]|\Phi_0\rangle = \omega_m \hat{R}(m)|\Phi_0\rangle, \tag{7}$$

$$\hat{R}(m) = \sum_{k=0}^{N} \hat{R}_k(m) = r_0(m) + \sum_{ai} r_i^a(m) a_a^\dagger a_i + \frac{1}{4} \sum_{abij} r_{ij}^{ab}(m) a_a^\dagger a_b^\dagger a_j a_i + \cdots, \tag{8}$$

$$\langle\Phi_0|\hat{L}(m)(\bar{H} - E_{CC}) = \langle\Phi_0|\hat{L}(m)\omega_m, \tag{9}$$

$$\hat{L}(m) = \sum_{k=1}^{N} \hat{L}_k(m) = \sum_{ai} l_a^i(m) a_i^\dagger a_a + \frac{1}{4} \sum_{abij} l_{ab}^{ij}(m) a_i^\dagger a_j^\dagger a_b a_a + \cdots, \tag{10}$$

Note that the left-hand eigenfunction equations are not explicitly connected. The connectedness of the right eigenfunction equations allows for a solution purely in terms of $\hat{R}_1$ and $\hat{R}_2$. Only one set of eigenfunctions is necessary to obtain the energy, although both are necessary in order to calculate properties (including transition properties) and analytic gradients.[26,27] We will use $\hat{R}(m)/\hat{L}(m)$ to refer to generic EOM-CC excitation/de-excitation operators, or a subscript EE to refer specifically to the EOMEE-CC amplitudes. STEOMEE-CC requires, in addition, singly ionized and electron-attached states which are formed via the application of non-number-conserving excitation operators (here only for the right-hand side),

$$\hat{R}_{IP}(m) = \sum_{k=1}^{N} \hat{R}_{IP;k}(m) = \sum_{i} r_i(m) a_i + \frac{1}{2} \sum_{aij} r_{ij}^{a}(m) a_a^\dagger a_j a_i + \cdots \tag{11}$$

$$\hat{R}_{EA}(m) = \sum_{k=1}^{N} \hat{R}_{EA;k}(m) = \sum_{a} r^{a}(m) a_a^\dagger + \frac{1}{2} \sum_{abi} r_{i}^{ab}(m) a_a^\dagger a_b^\dagger a_i + \cdots \tag{12}$$

2.2 STEOM-CC

Similarity transformed equation-of-motion coupled cluster theory[14] starts with the definition of the transformed Hamiltonian, $\hat{G}$,

$$\{e^{\hat{S}}\}\hat{G} = \bar{H}\{e^{\hat{S}}\}, \tag{13}$$

$$\hat{G} = \sum_{pq} g_q^p a_p^\dagger a_q + \frac{1}{4}\sum_{pqrs} g_{rs}^{pq} a_p^\dagger a_q^\dagger a_s a_r + \cdots \tag{14}$$

where the braces denote operator normal ordering (essentially, this ensures that $\{e^{\hat{S}}\}$ has no internal contractions). The transformation operator $\hat{S}$ has components in both the electron-attached, or $(1,0)$, and ionized, or $(0,1)$, sectors of Fock space,

$$\hat{S} = \hat{S}^+ + \hat{S}^-, \tag{15}$$

$$\hat{S}^+ = \sum_k \hat{S}_k^+ = \sum_{ae} s_e^a a_a^\dagger a_e + \frac{1}{2}\sum_{abej} s_{je}^{ba} a_a^\dagger a_b^\dagger a_j a_e + \cdots, \tag{16}$$

$$\hat{S}^- = \sum_k \hat{S}_k^- = \sum_{im} s_i^m a_m^\dagger a_i + \frac{1}{2}\sum_{ijmb} s_{ji}^{bm} a_m^\dagger a_b^\dagger a_j a_i + \cdots, \tag{17}$$

where e and m are associated with sets of $n_{v;act}$ active virtual and $n_{o;act}$ active occupied orbitals (transformed back to the canonical MO space).

As noted by Nooijen,[14] the $\hat{S}_1$ amplitudes are not necessary for the solution of the STEOMEE-CC equations as they simply cause rotations within the singles excitation space, and hence a diagonalization within the *full* singles space is invariant. In fact, diagonalization within the full rather than the active singles space is desirable as the portion of the solution falling outside the active space serves as a measure of active space insufficiency.[14]

The transformation amplitudes $\hat{S}^\pm$ may be easily obtained by renormalization of a set of

solutions of the EOMIP-CC and EOMEA-CC equations,

$$S_{ji}^{bm} = -\sum_{\kappa\lambda=1}^{n_{o;act}} r_{ji}^{b}(\lambda)\left(U_{-}^{-1}\right)_{\lambda\kappa}\delta_{\kappa m}, \tag{18}$$

$$S_{je}^{ba} = \sum_{\kappa\lambda=1}^{n_{v;act}} r_{j}^{ba}(\lambda)\left(U_{+}^{-1}\right)_{\lambda\kappa}\delta_{\kappa e}, \tag{19}$$

where the extra minus sign for the IP-coefficients comes from the contraction over the hole line λ in (18). The matrices $U_{\pm}$ are the transformation matrices which diagonalize the STEOM effective Hamiltonian. They are derived from the single excitation parts of the IP/EA solutions,

$$\left(U_{-}\right)_{\kappa\lambda} = \sum_{n}\delta_{\kappa n}r_{n}(\lambda), \tag{20}$$

$$\left(U_{+}\right)_{\kappa\lambda} = \sum_{f}\delta_{\kappa f}r^{f}(\lambda). \tag{21}$$

In (18)–(21), the factor $\delta_{\kappa p}$ indicates that the active orbitals (indexed by κ) are simply a subset of the canonical molecular orbitals (indexed by p), typically the orbitals within a small energy range around the Fermi level. The solutions $\hat{R}_{IP/EA}(\lambda)$ typically correspond to principal ionizations from and electron attachments to these active orbitals. The active orbitals may also be chosen as linear combinations of molecular orbitals,[23,28] with the proper transformation matrix replacing the Kronecker delta.

The excited states are obtained by solving the eigenvalue equations,

$$[\hat{G}, \hat{R}_1(m)]|\Phi_0\rangle = \omega_m \hat{R}_1(m)|\Phi_0\rangle. \tag{22}$$

The left-hand eigenvalue equations formally require a solution in the full singles and doubles space, although computation of properties and transition strengths can be simplified by a perturbative approximation of $\hat{L}_2$.[14]

2.3 Core Excited States

The direct calculation of core excited states, even for the 1s orbitals of first row elements with energies on the order of 100–500 eV, is fraught with difficulties. Standard EOMEE-CC calculations are difficult or impossible to converge, and even when convergence is achieved, the energies may be contaminated by spurious couplings to high-lying valence excited determinants which form an unphysical discretization of the valence continuum.[29]

As a remedy to this problem, Coriani and Koch[16] adapted the core-valence separation scheme first introduced by Cederbaum, Domcke, and Schirmer[30] to EOM-CC. In CVS-EOM-CC, amplitudes which correspond to purely valence excitations are explicitly zeroed. This leaves only components involving one or more core orbitals (indicated by capital roman letters),

$$\hat{R}_{CVS} = \sum_{aI} r_I^a(m) a_a^\dagger a_I + \frac{1}{2} \sum_{abIj} r_{Ij}^{ab}(m) a_a^\dagger a_b^\dagger a_j a_I + \frac{1}{4} \sum_{abIJ} r_{IJ}^{ab}(m) a_a^\dagger a_b^\dagger a_J a_I + \cdots . \tag{23}$$

This formulation eliminates all core-valence couplings and recovers the core-excited states as bound solutions. A number of variations on this basic theme exist, such as including only one or a small number of symmetry-related core orbitals in the "core" set and treating the rest as valence,[21] solving the ground state equations in the frozen-core approximation,[31] etc. In this work we target a single core orbital in each calculation and define $\hat{R}$ as in (23), with an all-electron solution of the ground state. Note that we purposefully avoid test molecules with symmetric nuclei as these cases require either treating the symmetric and anti-symmetric core molecular orbitals equally or breaking the symmetry via orbital localization.

Within STEOMEE-CCSD, the CVS can be applied in two places, giving rise to CVS-STEOMEE-CCSD. First, the selection of the active space must include the core orbital(s) of interest, and when solving for these core ionization wavefunctions the CVS is necessary to stabilize convergence. Second, the solution of the eigenvalue equations (22) may also use the CVS in defining the excitation operator $\hat{R}_1$ in order to accelerate the computation. Note

that if only one core orbital is included in the CVS treatment, then the diagonalization step scales as only $\mathcal{O}(n_{root}n_v^2)$, where n_v is the number of virtual orbitals. The application of the CVS to the $\hat{R}_{IP}$ operator does not induce a similar restriction on the $\hat{S}^-$ operator. Because this operator contains information from *both* valence and core ionized wavefunctions, it must necessarily span the full molecular orbital space. The combined valence and core nature of $\hat{S}^-$ also allows for the simultaneous and balanced determination of core and valence excitation energies.* In principle it may be possible to construct an $\hat{S}^-$ and hence $\hat{G}$ operator using *only* core ionization wavefunctions if no valence excitations are desired. We have not explored this possibility as the valence IP solutions are rarely a bottleneck in practice.

The CVS eliminates most convergence issues and recovers excitation energies which are systematically improvable towards the experimental value with increasing basis set size and level of excitation.[21,32] However, at the singles and doubles level, large errors remain which are attributable to the significant orbital relaxation from the ground to the excited core-hole state. We previously introduced the transition-potential coupled cluster method (TP-CC),[24] which accounts for orbital relaxation explicitly by performing the calculation with molecular orbitals optimized for a fractional core-hole. Instead, the STEOMEE-CC approach offers an alternative.

The ionization part of the similarity transformation, defined by the solution of the standard EOMIP-CC equations, provides all of the necessary dynamical correlation for the occupied orbitals. Then, since the inclusion of triple excitations in EOMEE-CC provides a necessary level of correlation to account for core-hole orbital relaxation, we propose that inclusion of triple excitations, *only in the EOMIP-CC solutions*, is sufficient to account for orbital relaxation in STEOMEE-CC. This modification introduces an additional transformation operator $\hat{S}_3^-$,

$$S_{kji}^{cbm} = -\sum_{\kappa\lambda=1}^{n_{O;act}} r_{kji}^{cb}(\lambda)\left(U_-^{-1}\right)_{\lambda\kappa}\delta_{\kappa m},$$

*The ground state, IP/EA, and $\hat{G}$ computations may be shared, although separate diagonalization steps are still necessary as the core and valence wavefunctions are not strictly orthogonal and have different structure. The diagonalization(s) are computational very inexpensive, however.

where at least one of ijk must be an active core orbital, and $\kappa\lambda$ refer to the $n_{O;act}$ active core orbitals. Since STEOMEE-CC only requires elements of $\hat{G}$ with at most two lines at the top and bottom, diagrammatically, the only modifications necessary to the formation of the twice-transformed Hamiltonian are,

$$g_i^m \leftarrow S_{kji}^{cbm} \langle jk \| bc \rangle,$$
$$g_{ei}^{ma} \leftarrow -S_{ijk}^{abm} \langle jk \| be \rangle.$$

The diagonalization step proceeds exactly the same as in unmodified CVS-STEOM-CCSD. We denote this new method as CVS-STEOMEE-CCSD+cT ("singles and doubles plus core triples"). For a single core orbital, the solution of the CVS-EOMIP-CCSDT equations scales as $\mathcal{O}(n_o^2 n_v^4)$, which is the same as the ground state CCSD equations. Also note that the connectivity of the EOMIP-CCSDT equations is preserved even without including triple excitations in the ground state.

Although we expect CVS-STEOMEE-CCSD+cT to be similar in overall accuracy to the existing CVS-EOM-CCSD* and TP-CCSD(1/2) methods, there are several advantages to CVS-STEOMEE-CCSD+cT over these alternatives. First, CVS-STEOMEE-CCSD+cT is more computationally efficient than CVS-EOM-CCSD* as triple excitations are only included in the ionization potential calculations, and not in the excitation energy calculation. In the case of CVS-EOM-CCSD*, the triple excitation contributions must also be included for each excitation, while for CVS-STEOMEE-CCSD+cT these contributions are only included once for each core orbital, irrespective of the number of excitations sought from that orbital. Second while we may expect CVS-STEOMEE-CCSD+cT and TP-CCSD(1/2) to incur similar overall computational costs, a TP-CCSD calculation is inherently specific to one core orbital/edge, as the reference orbitals must be specifically optimized. In CVS-STEOMEE-CCSD+cT the reference orbitals are the standard canonical orbitals, and importantly, the valence excitations obtained are almost precisely the same as within standard STEOM-CCSD

(except for a very small core orbital contribution). We expect this property to be important for the calculation of excited state x-ray spectra, e.g. for transient XAS spectroscopy. A final important advantage of STEOM-CC in general is that a large number of excitations can be calculated without significantly raising computational cost or memory consumption.

3 Computational Details

CVS-STEOMEE-CCSD and CVS-STEOMEE-CCSD+cT were implemented in a development version of the CFOUR program package.[33] In all cases we included all canonical orbitals with orbital energies between -20 and +10 eV as active orbitals. A single core orbital was included in the CVS treatment and STEOM principal IP solution in each calculation. We observed only very small (a few 10s of meV) changes upon expansion of the active space.

The test set consisted of four vertical core excitation energies from each 1s core orbital of H_2O, CO, HCN, HF, HOF, HNO, CH_2, CH_4, NH_3, H_3CF H_3COH, H_2CO, H_2CNH, and H_2NF. This includes a total of 94 vertical excitation energies. The core excitations were selected as those for which we could reliably converge all methods tested, which typically consisted of the first four excitations of dominant single excitation character. All calculations utilized the aug-cc-pCVTZ basis set with all electrons correlated, except for H_2O where aug-cc-pCVQZ was used. In order to avoid complications due to missing relativistic effects, basis set incompleteness (particularly for Rydberg core excitations), geometric effects, and data quality and availability, which would all be a concern when comparing directly to experimental data, we have used full CVS-EOM-CCSDT as a benchmark, as in previous work.[24]

4 Results and Discussion

In the following discussion and in Figs. 1 and 2, the "shortened" names of the CVS-EOM methods will be used, e.g. CCSD = CVS-EOM-CCSD, with the exception of TP-CCSD(1/2).

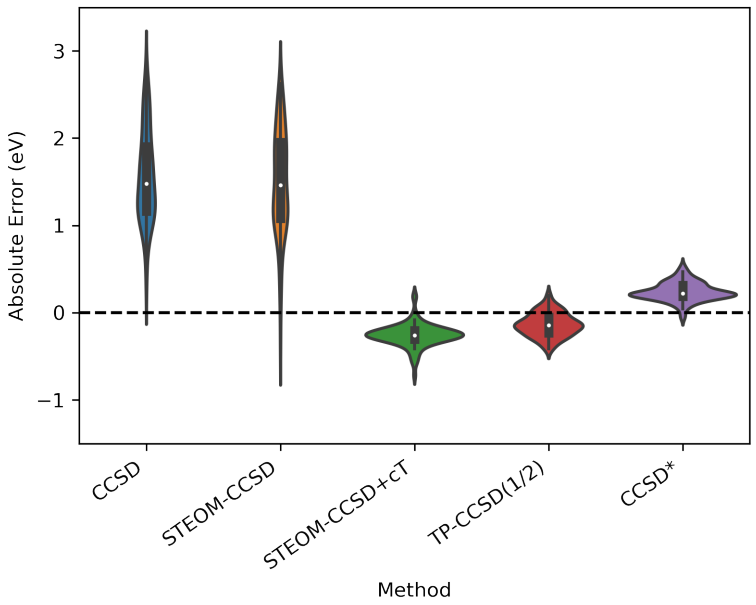
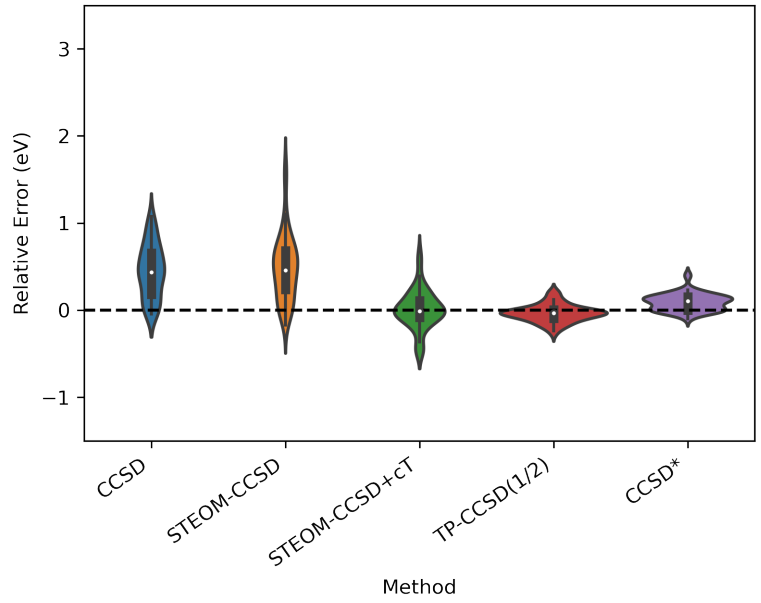


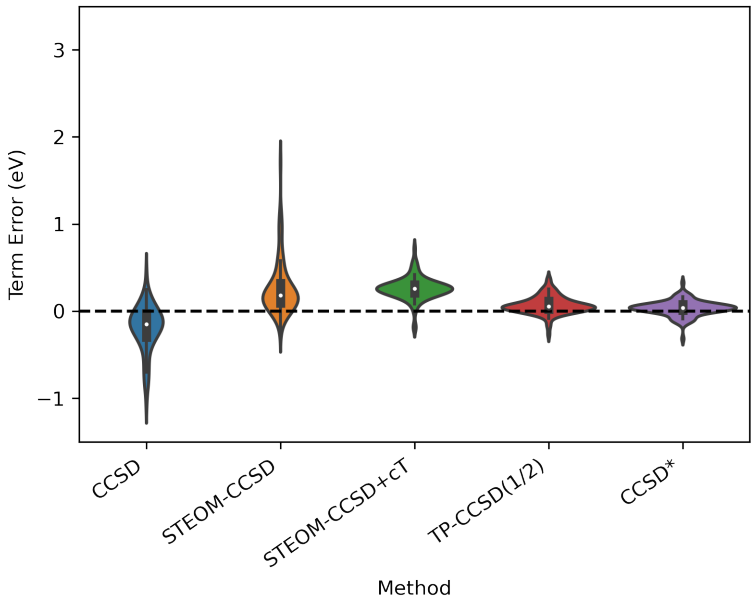
Figure 1: Error distributions with respect to CVS-EOM-CCSDT for absolute vertical core excitation energies.

The distribution of "absolute" (i.e. unmodified vertical) excitation energy deviations from CCSDT are depicted in Figure 1. The absolute energy deviation for a method X is calculated as $E(X) - E(\text{CCSDT})$ where E is a vertical core excitation energy. The "relative" excitation energy deviations are depicted in Figure 2. These deviations are determined in two ways: a) from excitation energies measured relative to the lowest excitation energy in each edge $E_{rel}(X) = E(X) - E_1(X)$, and b) from excitation energies measured downwards from the ionization edge, $E_{rel}(X) = IP(X) - E(X)$, and then deviations are computed as before. In the former case the first excitation energy (which has zero relative error by definition) is not included in the distribution. The use of the ionization edge or excitation onset as a reference causes a shift of the entire spectrum which is method- and molecule-specific. Note that the core ionization potentials for CVS-STEOMEE-CCSD are identical to EOMIP-CCSD, while those for CVS-STEOMEE-CCSD+cT are the same as "EOMIP-CCSD(2,3)", i.e. EOMIP-CC with a CCSD ground state and an EOMIP-CCSDT ionized state.

From Figure 1, it is clear that both "purely singles and doubles" methods, CCSD and STEOM-CCSD are prone to large errors in absolute vertical core excitation energies. In the



(a)



(b)

Figure 2: Error distributions for relative vertical excitation energy errors with respect to CVS-EOM-CCSDT. (a) Error in position relative to lowest excitation energy within the pre-edge region, and (b) error relative to the corresponding ionization edge.

case of CCSD, errors are as large as 3 eV but uniformly positive, reflecting the tendency of orbital relaxation to lower the excitation energies. STEOM-CCSD errors cover a slightly larger range, with a few states predicted too low, but overall the performance of the two methods is similar. The remaining methods: TP-CCSD(1/2) (which includes explicit orbital relaxation), CCSD* (which includes explicit triple excitations in the excited state), and our new method STEOM-CCSD+cT, all improve significantly, lowering absolute errors to less than 1 eV and less than 0.5 eV in the typical case.

On the other hand, Figure 2 represents a more useful picture of error characteristics by including a shift of the spectrum for each ionization edge. Since such as shift is almost always applied when comparing to experimental spectra, these error distributions reflect the remaining errors which directly affect the structure of the pre-edge region. Large errors here (larger than the experimental line widths of ~ 0.3 eV) can lead to errors in assignment and analysis of experimental spectra. Errors for CCSD and STEOM-CCSD are indeed reduced, although observed deviations from CCSDT cover almost 1.5 eV for CCSD and 2.5 eV for STEOM-CCSD. Here we note that STEOM-CCSD does indeed display noticeably worse performance compared to CCSD, with the larger number of outliers (almost all valence π^* states) significantly stretching the error distribution. This effect is particularly striking in Figure 2b where the otherwise nicely compact error distribution of STEOM-CCSD is ruined by large positive errors for such valence states (here, a positive error indicates that the valence states lie too low, or conversely that the Rydberg series lies too high). TP-CCSD(1/2) and CCSD* essentially achieve the goal of relative errors in the range of 0.3 eV for both types of relative error.

Errors for STEOM-CCSD+cT relative to the ionization edge (Figure 2b) show a fairly compact distribution but an overall downward shift of ~ 0.3 eV for the entire excitation spectrum. This may be due to imbalance in the level of correlation of the core ionization potential, which now includes full triples, and the excitation spectrum, which derives largely from the valence electron affinity calculation which remain purely singles and doubles. An

approximate treatment of triples in the core ionization potential calculation would further speed up the calculation and also possibly correct for this excessive gap. When comparing to the lowest excitation energy (Figure 2a), STEOM-CCSD+cT displays a compact distribution flanked by two outlying wings. The positive wing is essentially entirely due to CH_2, where the gap between the excitation into the lone pair and the Rydberg states is overestimated by 0.6 eV. The negative wing is dominated by the fluorine K-edges of HOF and H_2NF, which underestimate the gap between the mixed $\sigma^*/3s$ excitation and the remaining excitations by 0.4–0.5 eV. While these cases represent challenging electronic structures for STEOM-CCSD+cT, the overall pre-edge structure is maintained well except for these single gaps. Noticeably, the large errors for π^* valence states present in STEOM-CCSD are almost entirely eliminated in STEOM-CCSD+cT due to the improved description of the core hole.

5 Conclusions and Future Work

The problem of orbital relaxation is central to the accurate computation of core-hole spectra. While our previous work focused on an explicit inclusion of core relaxation via the use of orbitals optimized for fractional core occupation, here we show that a simple modification of similarity-transformed equation-of-motion theory can similarly address this important problem. Using full CVS-EOM-CCSDT as a benchmark, our new CVS-STEOMEE-CCSD+cT method reduces errors in absolute vertical core excitation energies by a factor of ~ 5 compared to CVS-EOMEE-CCSD, and also improves errors in pre-shifted excitation spectra to less than 0.5 eV, except for the challenging cases of CH_2 and fluorine K-edges with valence-Rydberg mixing. The addition of triple excitations only affects the core ionization calculation and scales as $\mathcal{O}(n_c n_o^2 n_v^4)$ for n_c active core, n_o occupied, and n_v virtual orbitals. This results in a modest increase in computational effort compared to standard CVS-EOM-CCSD, since the solution of the electron-attachment problem is typically the main bottleneck in STEOM-CC other than the ground state.

We are currently implementing CVS-STEOM-CC excited state and transition properties, which require little additional development over the standard theory.[14,23] Importantly, the inclusion of triple excitations in the core ionization potential calculation does not directly enter the computation of these properties. However, inclusion of triple excitations should indirectly improve the quality of the transition moments. We are currently working on implementing these properties. In another vein, the approximate inclusion of triple excitations may offer an equally effective and even more computationally inexpensive option. With the advances developed in this and future work, we are confident that similarity-transformed equation-of-motion theory will find significant use in core-hole spectroscopy, not only due to the excellent accuracy exhibited in our benchmark, but also due to the ability of STEOM to treat large numbers of excited states efficiently, its ability to treat core and valence states on an even footing, and its proven computational efficiency.

Acknowledgments

The authors would like to thank Prof. Marcel Nooijen for inspiring us to work on triple excitations in STEOM-CC theory and for several fruitful discussions. This work was supported in part by the US National Science Foundation under grant OAC-2003931. MS is supported by an SMU Center for Research Computing Graduate Fellowship. All calculations were performed on the ManeFrame II computing system at SMU.

Supplementary Material

An electronic supplementary information file is available as an Excel file (.xslx). This file contains the employed molecular geometries and basis sets and the complete results, comprising absolute and relative excitation energies, and ionization potentials.

Data Availability

The data that supports the findings of this study are available within the article and its supplementary material.

References

(1) Bartlett, R. J.; Musiał, M. Coupled-Cluster Theory in Quantum Chemistry. *Rev. Mod. Phys.* **2007**, *79*, 291–352.

(2) Mukherjee, D.; Mukherjee, P. K. A Response-Function Approach to the Direct Calculation of the Transition-Energy in a Multiple-Cluster Expansion Formalism. *Chemical Physics* **1979**, *39*, 325–335.

(3) Koch, H.; Jensen, H. J. A.; Jørgensen, P.; Helgaker, T. Excitation Energies from the Coupled Cluster Singles and Doubles Linear Response Function (CCSDLR). Applications to Be, CH+, CO, and H2O. *J. Chem. Phys.* **1990**, *93*, 3345.

(4) Geertsen, J.; Rittby, M.; Bartlett, R. J. The Equation-of-Motion Coupled-Cluster Method: Excitation Energies of Be and CO. *Chem. Phys. Lett.* **1989**, *164*, 57–62.

(5) Comeau, D. C.; Bartlett, R. J. The Equation-of-Motion Coupled-Cluster Method. Applications to Open- and Closed-Shell Reference States. *Chem. Phys. Lett.* **1993**, *207*, 414–423.

(6) Stanton, J. F.; Bartlett, R. J. The Equation of Motion Coupled-Cluster Method. A Systematic Biorthogonal Approach to Molecular Excitation Energies, Transition Probabilities, and Excited State Properties. *J. Chem. Phys.* **1993**, *98*, 7029.

(7) Nakatsuji, H.; Hirao, K. Cluster Expansion of the Wavefunction. Symmetry-adapted-cluster Expansion, Its Variational Determination, and Extension of Open-shell Orbital Theory. *J. Chem. Phys.* **1978**, *68*, 2053–2065.

(8) Nakatsuji, N. Cluster Expansion of the Wavefunction. Excited States. *Chemical Physics Letters* **1978**, *59*, 362–364.

(9) 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.

(10) Krylov, A. I. Spin-Flip Equation-of-Motion Coupled-Cluster Electronic Structure Method for a Description of Excited States, Bond Breaking, Diradicals, and Triradicals. *Acc. Chem. Res.* **2006**, *39*, 83–91.

(11) 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.

(12) Musiał, M.; Perera, A.; Bartlett, R. J. Multireference Coupled-Cluster Theory: The Easy Way. *J. Chem. Phys.* **2011**, *134*, 114108.

(13) Nooijen, M.; Bartlett, R. J. A New Method for Excited States: Similarity Transformed Equation-of-Motion Coupled-Cluster Theory. *J. Chem. Phys.* **1997**, *106*, 6441–6448.

(14) Nooijen, M.; Bartlett, R. J. Similarity Transformed Equation-of-Motion Coupled-Cluster Theory: Details, Examples, and Comparisons. *J. Chem. Phys.* **1997**, *107*, 6812–6830.

(15) Norman, P.; Dreuw, A. Simulating X-ray Spectroscopies and Calculating Core-Excited States of Molecules. *Chem. Rev.* **2018**, *118*, 7208–7248.

(16) 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.

(17) 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* **2012**, *85*, 022507.

(18) Zuev, D.; Jagau, T.-C.; Bravaya, K. B.; Epifanovsky, E.; Shao, Y.; Sundstrom, E.; Head-Gordon, M.; Krylov, A. I. Complex Absorbing Potentials within EOM-CC Family of Methods: Theory, Implementation, and Benchmarks. *J. Chem. Phys.* **2014**, *141*, 024102.

(19) Matz, F.; Jagau, T.-C. Molecular Auger decay rates from complex-variable coupled-cluster theory. *J. Chem. Phys.* **2022**, *156*, 114117.

(20) Skomorowski, W.; Krylov, A. I. Feshbach-Fano approach for calculation of Auger decay rates using equation-of-motion coupled-cluster wave functions. I. Theory and implementation. *J. Chem. Phys.* **2021**, *154*, 084124.

(21) Carbone, J. P.; Cheng, L.; Myhre, R. H.; Matthews, D.; Koch, H.; Coriani, S. *Advances in Quantum Chemistry*; Elsevier: United Kingdom, 2019; Vol. 79; pp 241–261.

(22) Matthews, D. A. EOM-CC Methods with Approximate Triple Excitations Applied to Core Excitation and Ionisation Energies. *Mol. Phys.* **2020**, *118*, e1771448.

(23) Ranga, S.; Dutta, A. K. A Core–Valence Separated Similarity Transformed EOM-CCSD Method for Core-Excitation Spectra. *J. Chem. Theory Comput.* **2021**, *17*, 7428–7446.

(24) Simons, M.; Matthews, D. A. Transition-Potential Coupled Cluster. *J. Chem. Phys.* **2021**, *154*, 014106.

(25) Shavitt, I.; Bartlett, R. J. *Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory*, 1st ed.; Cambridge University Press: Cambridge ; New York, 2009.

(26) 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.

(27) 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.

(28) Dutta, A. K.; Nooijen, M.; Neese, F.; Izsák, R. Automatic Active Space Selection for the Similarity Transformed Equations of Motion Coupled Cluster Method. *J. Chem. Phys.* **2017**, *146*, 074103.

(29) 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.

(30) Cederbaum, L. S.; Domcke, W.; Schirmer, J. Many-Body Theory of Core Holes. *Phys. Rev. A* **1980**, *22*, 206–222.

(31) Vidal, M. L.; Feng, X.; Epifanovsky, E.; Krylov, A. I.; Coriani, S. New and Efficient Equation-of-Motion Coupled-Cluster Framework for Core-Excited and Core-Ionized States. *J. Chem. Theory Comput.* **2019**, *15*, 3117–3133.

(32) Sarangi, R.; Vidal, M. L.; Coriani, S.; Krylov, A. I. On the Basis Set Selection for Calculations of Core-Level States: Different Strategies to Balance Cost and Accuracy. *Mol. Phys.* **2020**, *118*, e1769872.

(33) Matthews, D. A.; Cheng, L.; Harding, M. E.; Lipparini, F.; Stopkowicz, S.; Jagau, T.-C.; Szalay, P. G.; Gauss, J.; Stanton, J. F. Coupled-Cluster Techniques for Computational Chemistry: The CFOUR Program Package. *J. Chem. Phys.* **2020**, *152*, 214108.

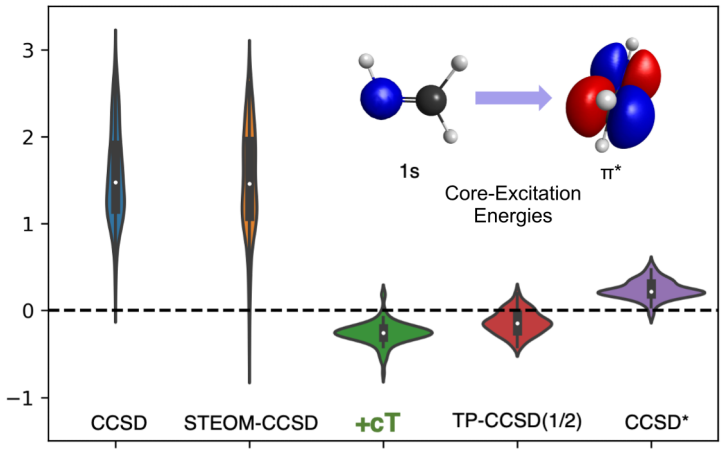


Figure 3: For Table of Contents Only