

UC Berkeley

UC Berkeley Previously Published Works

Title

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

Permalink

<https://escholarship.org/uc/item/2jz2t712>

Journal

Physical Chemistry Chemical Physics, 24(35)

ISSN

1463-9076

Authors

Arias-Martinez, Juan E
Cunha, Leonardo A
Oosterbaan, Katherine J
[et al.](#)

Publication Date

2022-09-01

DOI

10.1039/d2cp01998a

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

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

Juan E. Arias-Martinez,^{*,†,‡} Leonardo A. Cunha,^{†,‡} Katherine J. Oosterbaan,[†]
Joonho Lee,[¶] and Martin Head-Gordon^{*,†,‡}

[†]*Kenneth S. Pitzer Center for Theoretical Chemistry, Department of Chemistry, University of California, Berkeley, California 94720, USA*

[‡]*Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*

[¶]*Department of Chemistry, Columbia University, New York, New York 10027, USA*

E-mail: juanes@berkeley.edu; mhg@cchem.berkeley.edu

Abstract

We investigate the use of orbital-optimized references in conjunction with single-reference coupled-cluster theory with single and double substitutions (CCSD) for the study of core excitations and ionizations of 18 small organic molecules, without any use of response theory or equation-of-motion formalisms. Three schemes are employed to successfully address the convergence difficulties associated with the coupled-cluster equations, and the spin contamination resulting from the use of a spin symmetry-broken reference, in the case of excitations. In order to gauge the inherent potential of the methods studied, an effort is made to provide reasonable basis set limit estimates for the transition energies. Overall, we find that the two best-performing schemes studied

here for Δ CCSD are capable of predicting excitation and ionization energies with errors comparable to experimental accuracies. The proposed Δ CCSD schemes seem to fare better than the widely used equation-of-motion CCSD (EOM-CCSD) with core-valence separation protocol, with statistical errors being reduced by more than a factor of two when compared to FC-CVS-EOM-CCSD.

Introduction

Recent decades have seen improvements in the handling of high-energy ultra-violet (XUV) and X-ray light in terms of coherence,^{1,2} intensity,² and time control.²⁻⁴ As a result, scientists have been able to observe phenomena in chemistry,⁵⁻⁷ material sciences,^{8,9} and physics^{10,11} that were previously inaccessible. Furthermore, the increasing availability of table-top equipment¹²⁻¹⁵ capable of generating the light required for core spectroscopies has extended the use of said techniques for a variety of new studies.¹⁶ Efficiently and accurately modeling core excited states presents several challenges that a useful methodology should address, chief among them the large charge rearrangement associated with the creation of the core hole. Within the independent particle model, this charge rearrangement results in a strong contraction of the orbitals due to the decreased nuclear screening - this is referred to as orbital relaxation in the literature. The most widely used method for calculating valence excited states, time-dependent density functional theory (TD-DFT), struggles to describe core excited states (and charge transfer states in general) because the linear-response formalism fails to account for the charge rearrangement when standard exchange-correlation functionals are used.¹⁷⁻²⁰ Functionals that are custom-built for core excitations have, nonetheless, shown promise to bringing the computational cost advantages of DFT over to the realm of core spectroscopies.¹⁷ To circumvent the uncertainty associated with the choice of functionals, established wavefunction theories that are well-regarded for their accuracy in describing valence excitations, such as equation-of-motion coupled-cluster (EOM-CC) theory and algebraic diagrammatic construction (ADC), have been extended to core excitations by specifi-

cally targeting the very high-energy core excitation roots of their effective Hamiltonians.^{21–23}

An alternative approach followed by state-specific methods such as Δ SCF^{24,25} and its correlated relatives,^{26–33} the closely-related Transition Potential (TP)-SCF approaches,^{24,34–36} a number of multi-reference (MR) wavefunction models,^{37–40} excited state mean field theory,⁴¹ and Monte-Carlo-based approaches,⁴² is to account for relaxation in some way by optimizing for a target state. The Δ SCF approach, for example, converges a set of orbitals in a configuration that resembles the one-electron picture of the core excitation in question. These are non-Aufbau solutions to the self-consistent field (SCF) equations and are often saddle points in orbital space. Similarly, TP-SCF employs configurations optimized for a fractional core occupancy in the hopes of providing a reference of similar quality for both the ground and the core excited states. A difficulty of orbital-optimized excited state approaches is the possibility of landing on an undesired SCF solution of lower energy. In the context of mean-field approaches, such as Hartree-Fock (HF) and density functional theory (DFT), this issue has been addressed by algorithms specialized for excited state optimization, such as the maximum overlap method (MOM),⁴³ and, more recently, the initial MOM (IMOM),⁴⁴ square-gradient minimization (SGM)²⁵ and state-targeted energy projection (STEP)⁴⁵ methods.

Δ SCF has been used for decades to calculate core ionizations with success.^{24,26,29} In the cases where there are symmetry-equivalent atoms present in the system, an orbital localization procedure (such as that of Boys⁴⁶) must be carried out on the core orbitals prior to SCF re-optimization to allow for proper orbital relaxation.³⁰ The spatial symmetry breaking technically renders these situations multi-reference (MR) since multiple configurations must be re-combined via non-orthogonal configuration interaction (NOCI) to yield states of the proper spatial symmetry. In practice, the splitting between the symmetry-adapted configurations is small,^{47,48} so that the MR character associated with the core hole localization can be disregarded without serious error. The Δ SCF ionization energies, as calculated with the spatially symmetry-broken configurations are often good estimates of what would be observed in an experiment.

Studies on core excitations with Δ SCF have been more sparse until quite recently.²⁶ In some measure this is due to the fact that MR character now factors in because of the need for two configurations for a spin-pure description of the excited state. The approximate spin-projection scheme (AP) established a way to estimate the excitation energy of the pure singlet, provided that the energies of a spin-contaminated singlet and the pure triplet are known.^{49,50} An attractive alternative to AP for Δ SCF calculations is the use of restricted open-shell Kohn-Sham orbitals (ROKS), which optimizes the spin-pure singlet energy as computed via the AP scheme for a mixed and a triplet ($M_s = 1$) configuration sharing the same set of restricted open-shell (RO) orbitals.^{51–54} Recently, this technique (and a generalized version for radicals⁵⁵) has been used to study core excited states with the best-performing functional (SCAN) achieving an impressive 0.2 eV root-mean-squared-deviation (RMSD) from experimental results for a representative set of small organic molecules.^{25,56} With an appropriate treatment of scalar relativistic effects, ROKS has also been employed to tackle the K-edge of third-group elements.⁵⁷

Excited SCF solutions are often a better reference than the ground state for finding alternative solutions to the CC equations, which in turn are reasonable approximations to the true excited states.⁵⁸ Explicit SCF re-optimization takes care of the strong orbital relaxation and allows single-reference (SR) post-HF methods such as second order Møller-Plesset perturbation theory (MP2) and CC to focus on addressing the remaining dynamic correlation of a system. Core ionized states of closed-shell systems are perfect cases to be treated by these models and they have been studied via Δ MP2^{24,26–29} and, more recently, Δ CCSD(T).^{30–32} The last decade has seen an effort to also employ explicitly-relaxed orbitals on a (wavefunction-based) correlated calculation for singlet excited states.^{26,32,33,37–40,59} Among these, the wavefunction theories employing explicit MR construction often constrain them to study few molecules in small basis sets, which means they can only be compared to other computational methods in the same small basis sets.^{37–40} Simons and Matthews have recently proposed a theory, TP CC, that employs a TP SCF reference for an EOM-CC

calculation of the core excited states.⁶⁰ This model inherits some of the advantages of both state-specific methods - orbital relaxation - while retaining the advantages of EOM-CC: inherent spin-adaptation of the excited states, a full spectrum with a single calculation, and straightforward transition properties. The cost to pay comes from relying on a deteriorated description of the ground state relative to standard CC, controlled by tuning the fractional occupation number of the core orbital. Even though this renders the model arbitrary, to some extent, Simons and Matthews have carried out a study to find an optimal core occupancy parameter transferable across edges of the same element, making this a promising method for reliable and affordable high-accuracy wavefunction X-ray calculations.⁶¹

Owing to the simple nature of the MR character of singly core excited states of closed shell systems (namely, a two-determinant CSF) the objective of this paper is to assess the use of SR CC formalism with orbital-optimized references, limited to the level of single and double substitutions (i.e. CCSD) for computational tractability) for core excitations. In contrast to TP-CC, the protocols presented here are well-defined in that only the molecule and the transition of interest need to be specified - the ground state CC wavefunction and energies are used as is and no compromise in the excited state wave function is made either. The major issue to be addressed in order to permit such calculations is how to treat the electron correlation effects that couple core orbitals with either other core levels or valence levels. De-excitation into the core hole can lead either to numerical instabilities or variational collapse towards the ground state. Therefore a suitable adaptation of SR CCSD for state-specific optimization of core excited states must treat core correlation, as well as removing potentially ill-behaved amplitudes. Additionally, a correction to produce approximately spin-pure eigenstates is required.

This paper is organized as follows. After a review of the appropriate theory, we describe three candidate approaches that we deem potentially promising. Two of them employ Yamaguchi's AP approach,⁴⁹ while the third one instead enforces correct spin symmetry at the ROHF level by constraining the amplitude of the double substitution that flips the spins of

the two half-occupied orbitals to +1 for singlet and -1 for triplet states. A comparison of these approaches is then made against successful core-excited state theories, ROKS(SCF) and FC-CVS-EOM-CC, with the ultimate judge being the experimental results. The energetic differences between the singlet and triplet core excited states, presumed to be accurate enough to make a statement about them, are presented. An effort is made to reach basis set convergence for all methods in order to exclude this factor from the discussion as much as possible and focus on their inherent performance. Despite the computational demands of approaching the basis set limit (BSL) for CC methods constraining us to molecules with at most two heavy atoms, the data set is diverse in terms of the elements targeted (Be, C, N, O, F, Ne) and in terms of the excited state character (σ^* , π^* , Rydberg). In total, a set of 21 excitations and 18 ionizations on 18 small closed-shell organic molecules is used.

Background

Following convention, we will reserve the indexes $i, j, k \dots$ for any occupied orbital, $a, b, c \dots$ for any virtual orbital, and $p, q, r \dots$ for an arbitrary orbital. For the CCSD amplitudes, we will use the symbols t_i^a and t_{ij}^{ab} , collected in T_1 and T_2 .

For a set of orbitals that are not necessarily canonical, the CCSD amplitude equations take the following form:

$$D_i^{at^a} = F_{ia} + w_i^a(T_1, T_2) \tag{1}$$

$$D_{ij}^{abt_{ij}^{ab}} = \langle ij | ab \rangle + w_{ij}^{ab}(T_1, T_2) \tag{2}$$

The terms $w_i^a(T_1, T_2)$ and $w_{ij}^{ab}(T_1, T_2)$ in Eqs. 1 and 2 contain terms that are linear and higher in T_1 and T_2 separate from the orbital energy differences, D_i^a and D_{ij}^{ab} defined below.⁶²

$$D_i^a = \varepsilon_i - \varepsilon_a \tag{3}$$

$$D_{ij}^{ab} = \varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b \tag{4}$$

ε_p are the orbital energies themselves. D_i^a and D_{ij}^{ab} will always be negative when employing a ground state reference and, in the absence of strong correlation, are large enough to make the T amplitudes well behaved (i.e. $\max[|t_i^a|, |t_{ij}^{ab}|] \ll 1$).

State-specific optimization of a core excited state, on the other hand, correlates a non-Aufbau SCF reference. Here, we make use of three different kinds of such (beta) core-excited references: (i) open-shell, symmetry-broken $M_S = 0$ references for the calculation of the singlet core-excited states; open-shell, (ii) spin-pure triplet $M_S = 1$ references for the AP approach, when needed; and (iii) open-shell, spin-pure $M_S = \frac{1}{2}$ doublet references for the calculation of core ionized states. In the case of the spin-pure triplet and pure doublet references, standard ROHF is used in conjunction with MOM. The use of unrestricted orbitals for the symmetry-broken reference was found to be detrimental to some of our Δ CC schemes, so ROKS(HF) orbitals, followed by a Fock-build for the broken-symmetry singlet state and further pseudocanonicalization, were employed instead.

With these choices of reference, and specific to the case of core excitations, the presence of a virtual orbital with a large negative energy representing the core hole (we reserve the indexes h and \bar{h} for the occupied alpha core orbital and the virtual beta core orbital) allows for denominators D_i^a and D_{ij}^{ab} to be positive when $a = \bar{h}$. In the case of single excitations, $a_h^\dagger a_i$, this occurs when the occupied orbital has a higher orbital energy than the core virtual

$$\varepsilon_i > \varepsilon_{\bar{h}} \quad (5)$$

The condition in Eq. 5 holds unless there are other core orbitals of lower orbital energy. In the case of double excitations, $a_h^\dagger a_i a_b^\dagger a_j$, $D_{ij}^{\bar{h}b}$ will be positive when

$$\varepsilon_i + \varepsilon_j - \varepsilon_b > \varepsilon_{\bar{h}} \quad (6)$$

One scenario where this happens is when the excitation $a_h^\dagger a_i$ involves a valence occupied orbital and the excitation $a_b^\dagger a_j$ involves only valence orbitals. The denominator $D_{ij}^{\bar{h}b}$ can still

be negative if the other virtual has an orbital energy ε_b positive and large enough to break Eq. 6. Furthermore, the orbital energies can conspire to make $\varepsilon_i + \varepsilon_j - \varepsilon_b \approx \varepsilon_{\bar{h}}$, rendering $D_{ij}^{ab} \approx 0$. Depending on the ability of the basis set to describe the high-lying virtual orbitals associated with the continuum, the denominator associated with double excitations can get arbitrarily close to zero, leading to numerical difficulties in solving for the T amplitudes (and of course divergence of perturbation theory methods, such as MP2).

Close-to-zero denominators also yield numerical instabilities in the context of EOM-CC. In their study of EOM-CC-IP for K-edge ionization energies, Liu *et al.* found that spurious high-lying valence excited states that are quasi-degenerate with the core excited state result in erratically-converging correlation energies with respect to basis set.⁴⁷ The core-valence separation (CVS) scheme is a proposed solution to this numerical problem; in this approach, core excitations are excluded from the ground state amplitudes, and all valence excitations are excluded from the EOM amplitudes.²³ The spurious couplings with the high-lying continuum excited states are then removed by design.

In a spirit similar to the CVS scheme, Zheng *et al.* proposed to exclude the virtual core orbital from the correlation treatment to address the divergence problem in the Δ CC calculations of core ionizations.^{30,31} Some of us adopted a similar strategy where we freeze up the doubly-vacant core orbital all together when studying double-core excitations.³² Zheng *et al.* found the missing correlation to be relevant for accurate core ionizations and uses estimates from fully-correlated CC calculations with decreasing denominator thresholds to account for it.

Computational details

A development version of Q-Chem 5.4⁶³ was used for all calculations. Experimental geometries available on the NIST computational database⁶⁴ were used throughout this work. An atomic relativistic correction calculated via the Douglas-Kroll-Hell method, found to be

nearly independent of basis-set and molecule for the main group elements, is added to all calculations (0.012, 0.09, 0.18, 0.34, 0.57, and 0.91 eV for Be, C, N, O, F, and Ne.)⁶⁵ For two of the three schemes of Δ CC we employ, the calculated singlet excited states are spin contaminated; the AP method is used to estimate the spin-pure excitation energies.

Our best attempt was made at comparing the excitation or ionization energies near their BSL values. To that end, different procedures involving specialized basis sets were employed for obtaining an approximate BSL for the different methods. The aug-pcX-3 (heavy)/ aug-pcseg-2 (hydrogen) basis was used to approximate the BSL for the ROKS(SCF) calculations.⁶⁶ A (99, 590) Euler-Maclaurin-Lebedev grid was used for the computation of the exchange-correlation integrals for the ROKS(SCAN) calculations. The aug-ccX-nZ (heavy) / aug-cc-pVTZ (hydrogen) bases,⁶⁷ extrapolated using the two-point X^{-3} scheme^{68,69} with $n = T, Q$, were used to approximate the BSL for the EOM-CC calculations. As noted in a recent study, such an extrapolation scheme is appropriate for core excitations via EOM-CC.⁷⁰ All ROKS(SCF) and EOM-CC calculations were also run with the standard Dunning aug-cc-pCVXZ ($X = D, T, Q$) family of bases^{71,72} and a slower convergence towards a similar BSL value was observed (SI).

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

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

Approaches to inclusion of core-valence correlation

Scheme 0: Using the full set of amplitudes

To motivate the need for the schemes presented in the following subsections, we begin by exploring the behavior of the correlated methods with no modifications. The Fock matrix and MO coefficients of the optimized excited reference are passed to the correlated calculation and all amplitudes (e.g. all singles and doubles in CCSD) are included; we refer to this as Scheme 0 (S0). Scheme S0 would not be of use for real applications because of the possibility of variational collapse, and limitations of today’s standard iterative CC solvers. Nevertheless, it provides useful insight in the few cases where the coupled cluster equations do converge. Such systems are few-atom molecules in a small basis, where there are no orbitals of the right energy to make the denominators small enough.

Figure 1 shows the basis set convergence of the CH_4 core ionization energies, as calculated with the Δ -based methods, with respect to increasing cardinality of the aug-cc-pCVXZ basis set. The Δ SCF values converge quickly, with the 5Z result decreasing the calculated ionization energy by only 0.014 eV from the QZ numbers. The results for all the correlated Δ methods are within 0.1 eV of each other up until the QZ level, where they begin to diverge. At the 5Z level, the CCSD equations fail to converge and the Δ MP2 results break monotonicity. An analysis of the denominators associated with excitations into the core virtual (Figure 2) reveals that, for all basis sets, there are positive denominators and, furthermore, that a close-to-zero denominator appears at the QZ level. Once the complexity of the molecule increases, the virtual space will begin to populate the problematic orbital energy range associated with near-zero denominators even when using small basis sets. Yet the CCSD(S0) results, at the very least, suggest that accurate results via Δ -based methods could be obtained if the irregularities caused by small denominators were addressed.

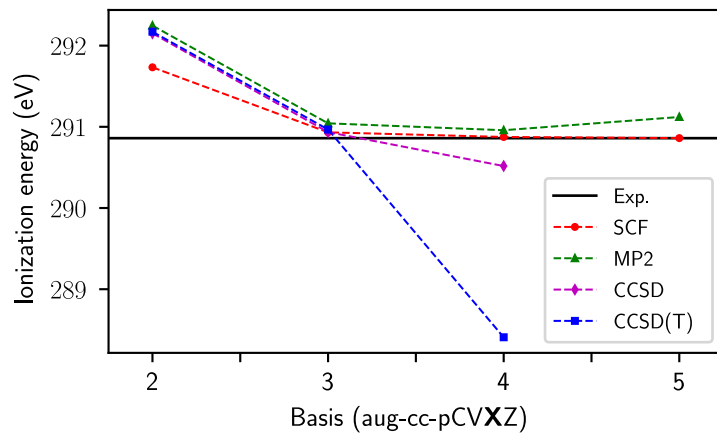


Figure 1: CH₄ ionization at the Frank-Condon geometry

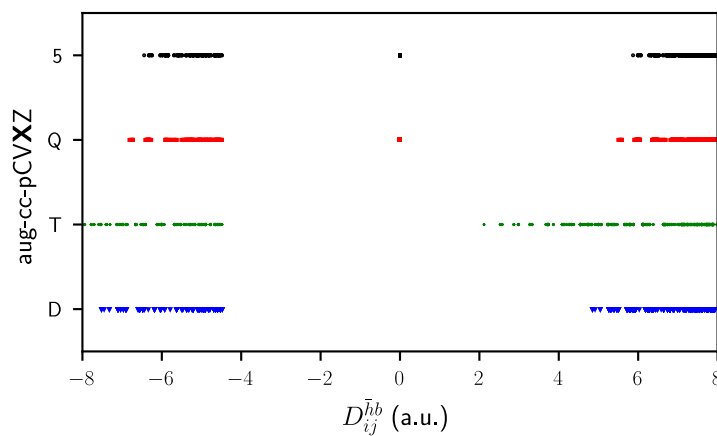


Figure 2: Some values for the denominators associated with excitation into the core virtual for the CH₄ core-ionized reference.

Scheme 1: Deleting all amplitudes involving the core virtual

We make use of three additional schemes to address the numerical instabilities discussed previously. The first, which we refer as Scheme 1 (S1), is that proposed Zheng et al.,³⁰ and employed by Lee and Head-Gordon.³² This scheme simply excludes any amplitude involving the core virtual. Additionally, we chose to exclude singles amplitudes that excite the occupied core electron.

$$\text{if}(a = \bar{h} \text{ or } i = h) \quad a_i^a, t_i^a = 0$$

$$\text{if}(a = \bar{h} \text{ or } b = \bar{h}) \quad a_{ij}^{ab}, t_{ij}^{ab} = 0$$

$$\text{if}(a = \bar{h} \text{ or } b = \bar{h} \text{ or } c = \bar{h}) \quad t_{ijk}^{abc}(c) = 0$$

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

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

To incorporate some of the correlation missing in S1, Scheme 2 (S2) allows for the double substitutions involving the core virtual, \bar{h} , that also promote the occupied electron in the same core orbital, h - these were found to be the leading amplitudes for some of the larger well-behaved S0 calculations. S2 is pleasing in that, even though core substitutions are

involved, they are all associated with configurations that retain a core occupancy of 1.

$$\text{if}(a = \bar{h} \text{ or } i = h) \quad a_i^a, t_i^a = 0$$

$$\text{if}(a = \bar{h} \text{ or } b = \bar{h})$$

$$\text{if}(i \neq h \text{ or } j \neq h) \quad a_{ij}^{ab}, t_{ij}^{ab} = 0$$

$$\text{if}(a = \bar{h} \text{ or } b = \bar{h} \text{ or } c = \bar{h})$$

$$\text{if}(i \neq h \text{ or } j \neq h \text{ or } k \neq h) \quad t_{ijk}^{abc}(c) = 0$$

As for S1, the CC de-excitation amplitudes are treated in a completely analogous way. We found that, in the case of the mixed singlets, allowing for the double substitution that generates the spin complement of the reference, $a_{\bar{h}}^\dagger a_{\bar{t}} a_t^\dagger a_h$ with t being the target orbital, leads the CC iterations to converge towards the (lower energy) triplet excited state, resulting $\langle S^2 \rangle$ values that deviate significantly from 1. Therefore, an additional constraint was placed on calculations for the mixed singlet: the amplitude associated with said excitation is also set to zero. This helped ensure that the $\langle S^2 \rangle$ value of the CCSD wave function remained close to 1, signifying that it is a mixed spin configuration. Therefore, as with S0, the spin contamination is removed by evaluating the singlet energy via Yamaguchi's AP expression.

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

As a final scheme, and exclusively for the calculations on the mixed singlet state, we propose to incorporate all of the conditions of S2 but, instead of neglecting the double substitution amplitude, $a_{\bar{h}}^\dagger a_{\bar{t}} a_t^\dagger a_h$, associated with the spin complement of the reference, we set it to -1.0; we refer to this as Scheme 3 (S3). These conditions force the CC iterations to look for the pure singlet starting from the mixed reference. As previously, the exact same S3 conditions are imposed on the de-excitation amplitudes for the left eigenvectors of the similarity trans-

formed Hamiltonian. We found that the lambda equations were able to converge even when the de-excitation amplitude associated with the spin complement is not forced to be -1.0. Enforcing said condition accelerated the convergence to result in the same value for $\langle S^2 \rangle$. An attractive feature of S3, as will be elaborated on in the Results section, is that it bypasses the need for AP altogether because the resulting states have $\langle S^2 \rangle$ values relatively close to 0. S3 is, in fact, similar in spirit to the the bi-configurational MR-CC model proposed by Oliphant and Adamowicz in 1991,⁷³ (see also the two-determinant Hilbert-space MR-CC,^{74,75} recently employed by Matthews⁵⁹ in conjunction with an orbital-optimized CSF for core excited states). However S3 is dramatically simpler because additional triple and quadruple excitations that are necessary in MR-CC (in order to account for the single and double excitations on top of the “secondary reference”) are omitted here.

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

Results and discussion

Before discussing the correlated methods, it is worth revisiting the spin-pure open shell singlet HF results (labeled as ROKS(HF) as this can be viewed as a special case of OO-DFT⁵⁴). For the excitations considered, ROKS(HF) achieves a mean absolute error (MAE) and RMSE of 0.43 and 0.52 eV. All of the excitations involving carbon and nitrogen, and the O 1s - σ^* /Rydberg transitions are overestimated. All of the fluorine and neon excitations, and the O 1s - π^* transitions are underestimated. This element-dependent error distribution with

respect to experiment leads to a relatively small mean signed error (MSE) of 0.18 eV. Using ROKS with the standard SCAN functional,⁷⁶ the best-performing functional according to a recent study, reduces the MAE to 0.16 eV, the RMSE to an impressive 0.19 eV,⁵⁶ and an MSE of only -0.08 eV. How well can CC methods limited to double or perturbative triple substitutions compete with these results?

As a further preliminary, we note that standard FC-CVS-EOM-CCSD approach cannot match ROKS(SCAN), and in fact scarcely outperforms the simple ROKS(HF) approach: EOM-CCSD achieves an MAE and RMSE of 0.34 and 0.41 eV. EOM-CCSD tends to underestimate the excitations out of carbon, with an overestimation of 0.34 eV for the $\text{CH}_3\text{OH } 1s \rightarrow 3s$ transition being the only serious exception. All other excitations are overestimated by EOM-CCSD, except for the $\text{N}_2 1s \rightarrow \pi^*$ and $\text{Be } 1s \rightarrow 2p$ excitations, which are underestimated by 0.25 and 0.68 eV, respectively. The latter might be a failure of the FC-CVS model.

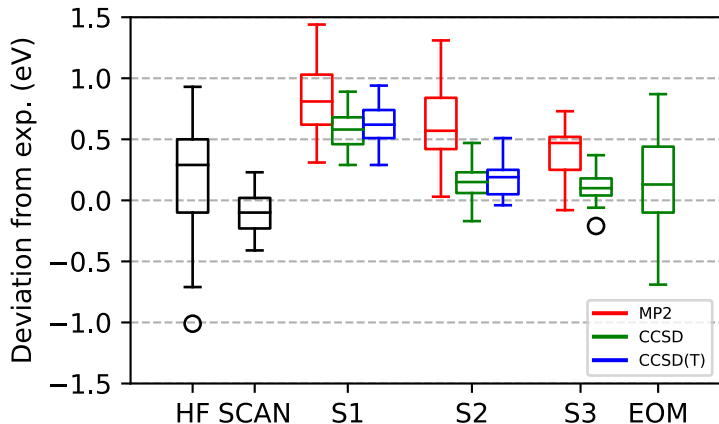


Figure 3: Statistical summary of the accuracy of calculated K-shell core excitations relative to experimental values for the 21 transitions shown in Table 1, as evaluated by ROKS(HF), ROKS(SCAN), the correlated Δ methods (Schemes S1, S2 and S3), and FC-CVS-EOM-CCSD-EE. For the S1 and S2 approaches, in addition to CCSD itself, the corresponding MP2 and CCSD(T) values are also shown. The specific values corresponding to these statistics are given in Table 1 and the Supplementary Information.

In regards to the correlated Δ methods, addressing the offending denominators, either by eliminating all excitations into the core virtual (S1) or including only those that retain a

core occupancy of 1 (S2 and S3) resulted in well-behaved, monotonically convergent CC calculations in all cases. Furthermore, for Schemes S1 and S2, the MP2, CCSD, and CCSD(T) correlation energies of the excited states, and the calculated excitation energies seem to converge monotonically towards a well defined BSL.

As observed in Figure 3 and the SI, correlated calculations via Scheme S1 always overestimate the excitation energy. $\Delta\text{MP2}(\text{S1})$, $\Delta\text{CCSD}(\text{S1})$, and $\Delta\text{CCSD}(\text{T})(\text{S1})$ achieve MAEs of 0.82, 0.58, 0.63 eV, and RMSEs of 0.88, 0.60, 0.65 eV. $\Delta\text{CCSD}(\text{S1})$ attenuates the most severe failures of $\Delta\text{MP2}(\text{S1})$ - where it overestimates experiment by more than 1 eV: $\text{H}_2\text{CO } 1s \rightarrow \pi^*$, $\text{HCN } 1s \rightarrow \pi^*$, $\text{HCN } 1s \rightarrow \pi^*$, $\text{N}_2 1s \rightarrow \pi^*$, and $\text{F}_2 1s \rightarrow \sigma^*$. It is worth noticing that these are all cases where $\Delta\text{MP2}(\text{S1})$ changes the ROKS(HF) results the most - in all cases for worse - with F_2 having the largest change of 2.3 eV. $\Delta\text{CCSD}(\text{T})(\text{S1})$, more often than not, seems to very slightly increase the error against experiment when compared to $\Delta\text{CCSD}(\text{S1})$. Including correlation via S1, either via MP2, CCSD, or CCSD(T) only decreases the calculated values relative to ΔHF in roughly half the cases. The MSEs for all the correlated methods under S1 are identical to their MAEs, which is consistent with a systematic overestimation of the excitation energies or, conversely, an under-correlation of the excited states. Since the results are expected to be well near the BSL, and the perturbative triples correction changes the CCSD results by a small amount, we attribute this to the configurations excluded from the correlation treatment for the sake of proper convergence.

Figure 3 and SI show that including some of the missing configurations via model S2 indeed reduces the error relative to S1. $\Delta\text{MP2}(\text{S2})$, $\Delta\text{CCSD}(\text{S2})$, and $\Delta\text{CCSD}(\text{T})(\text{S2})$ achieves MAEs of 0.62, 0.18, and 0.20 eV, and RMSEs of 0.69, 0.22 and 0.25 eV. A small systematic overestimation remains, as suggested by MSEs of 0.61, 0.16, and 0.20 eV. Two relevant statistical observations are that $\Delta\text{MP2}(\text{S2})$ still fails to offer an improvement over ROKS(HF), and that the (T) correction slightly worsens the ΔCCSD results. We note how the well-behaved excitations involving the core account for roughly 0.4 eV of the calculated excitation energy, as measured by the statistical differences between $\Delta\text{CCSD}(\text{S1})$ and $\Delta\text{CCSD}(\text{S2})$. This is in

agreement with the findings of Zheng *et. al*³⁰ and emphasizes that, if quantitative agreement is desired, a CVS scheme like S1 is inadequate.

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

The spin-forbidden excitations into the triplet $M_s = 0$ manifold were calculated with Δ CCSD(S3) by forcing the amplitude of the spin complement of the reference to be 1.0; they are listed in SI. We compared these against the triplet $M_s = 1$ excitation energies as calculated by Δ CCSD(S2). The largest deviation was of 0.09 eV for the H_2CO $1s \rightarrow \pi^*$ state, the average being 0.04 eV. The $M_s = 0$ triplet excitations were higher than the $M_s = 1$ results for all but one case, Be $1s \rightarrow 2p$, where the difference is -0.01 eV. This is also consistent with the idea that for the $M_s = 0$ triplets, as for the singlets, we are undercorrelating the excited state due to missing excitations. An undercorrelation is not present for the $M_s = 1$ triplet because, aside from any spatial symmetry breaking, this is purely a SR situation that S2 should be able to address. The triplet numbers, as calculated by Δ CCSD(S2), match fairly well with the two experimental numbers that we found for these spin-forbidden transitions: 114.3 eV for Be $1s \rightarrow 2p$ and 400.12 eV for N_2 $1s \rightarrow \pi^*$.⁷⁷ Δ CCSD(S2) predicts them to be 114.37 eV and 400.24 eV, respectively. The average energy difference between the singlet and triplet excited states for the set of molecules studied here, as calculated by Δ CCSD(S3), is 0.44 eV. Some cases worthy of notice are Be $1s \rightarrow 2p$

, where the splitting is 1.16 eV, and $\text{CO } 1s \rightarrow \pi^*$, with the largest splitting of all: 1.42 eV. Interestingly, the splitting for $\text{CO } 1s \rightarrow \pi^*$ is only 0.34 eV. Another case of relevance are the two Rydberg excitations $\text{Ne } 1s \rightarrow 3s$ and $\text{Ne } 1s \rightarrow 3p$, which have the smallest splittings: 0.06 eV and 0.05 eV, respectively.

In Table 1, we present the calculated excitation energies of the singlet excited states for the most successful scheme, $\Delta\text{CCSD}(\text{S3})$, against $\text{ROKS}(\text{HF})$, $\text{ROKS}(\text{SCAN})$, and $\text{FC-CVS-EOM-CCSD-EE}$.²³ All the statistics provided are compared against their measured experimental values. The per-molecule results for the remaining schemes are listed in the SI. Overall, $\Delta\text{CCSD}(\text{S3})$ achieves an MAE and RMSE of 0.14 and 0.18 eV. The most challenging excitation for this method is $\text{H}_2\text{CO } 1s \rightarrow \pi^*$, with an overestimation of 0.37 eV from the experimental value of 287.98 eV by Remmers *et al.*⁷⁸ A small systematic overestimation remains, as suggested by a MSE of 0.12 eV. The only excitation that $\Delta\text{CCSD}(\text{S3})$ significantly underestimates is $\text{CO } 1s \rightarrow \pi^*$, which is below Sodhi and Brion’s result of 534.21 ± 0.09 eV by 0.21 eV.⁷⁹ The closely-related two-determinant CCSD results of Matthews suggest a comparable accuracy, where a MAE of 0.10 eV and RMSE of 0.11eV were found against EOM-CCSDT-EE numbers for the three lowest lying core excitations of HCN, CO, NH_3 , and H_2O .⁵⁹

Table 2 compares the $\Delta\text{CCSD}(\text{S2})$ core ionizations, against those calculated by $\Delta\text{SCF}(\text{HF})$, $\Delta\text{SCF}(\text{SCAN})$ and $\text{FC-CVS-EOM-CCSD-IP}$. Figure 4 shows box-whisker plots for both the S1 and S2 methods applied to MP2, CCSD, and CCSD(T) relative to the same existing methods. The experimental values used as a reference are the ones given by Jolly *et al.*,⁸⁷ unless a more recent study was found. $\Delta\text{SCF}(\text{HF})$ has a MSE, MAE, and RMSE of -0.15, 0.45, and 0.58 eV, respectively. The two most challenging cases for ΔSCF in the ionization data set, CO and F_2 , are the only cases with an error greater than 1 eV. $\Delta\text{SCF}(\text{SCAN})$ reduces the $\Delta\text{SCF}(\text{HF})$ errors by more than a factor of two, with an MAE and RMSE of 0.21 and 0.25 eV. In contrast to excitations, all ionizations except two, F_2 and Ne, are overestimated with $\Delta\text{SCF}(\text{SCAN})$, resulting in an MSE similar to its MAE: 0.18 eV. The most

Table 1: Core excitations

Transition	ROKS(HF)	ROKS(SCAN)	Δ CCSD(S3)	EOM-CCSD	Experiment
Be 1s - 2p	115.37	115.34	115.53	114.79	115.47 ⁸⁰
C ₂ H ₄ 1s - π^*	285.27	284.70	284.77	284.68	284.68 (0.1) ⁸¹
H ₂ CO 1s - π^*	286.42	285.74	285.96	285.62	285.59 ⁷⁸
C ₂ H ₂ 1s - π^*	286.40	285.67	285.84	285.55	285.9 (0.1) ⁸¹
HCN 1s - π^*	286.98	286.35	286.51	286.07	286.37 ⁸²
CO 1s - π^*	288.05	286.99	287.46	286.71	287.40(0.02) ⁷⁹
CH ₃ OH C 1s - 3s	288.91	288.18	288.34	288.26	287.98 ⁸³
CH ₄ 1s - 3p(<i>t</i> ₂)	288.38	287.96	288.02	287.9	288.00 (0.2) ⁸⁴
HCN 1s - π^*	400.00	399.60	399.80	399.74	399.7 ⁸²
NH ₃ 1s - 3s	400.97	400.42	400.63	400.82	400.66 (0.2) ⁸⁴
N ₂ 1s - π^*	401.18	400.80	401.02	400.63	400.88 (0.02) ⁷⁹
NH ₃ 1s - 3p(e)	402.62	402.18	402.41	402.46	402.33 (0.2) ⁸⁴
H ₂ CO 1s - π^*	530.67	530.83	530.86	531.26	530.82 ⁷⁸
H ₂ O 1s - 3s	534.15	533.84	534.14	534.44	534.0 (0.2) ⁸⁴
CH ₃ OH 1s - 3s	534.16	533.98	534.24	534.64	534.12 ⁸³
CO 1s - π^*	533.68	533.97	534.00	534.50	534.21 (0.09) ⁷⁹
H ₂ O 1s - 3p (<i>b</i> ₂)	536.03	535.65	536.08	536.21	535.9 (0.2) ⁸⁴
F ₂ 1s - σ^*	681.19	682.43	682.41	683.07	682.2 (0.1) ⁸⁵
HF 1s - σ^*	687.31	687.44	687.76	688.05	687.4 (0.2) ⁸⁵
Ne 1s - 3s	864.75	865.18	865.37	865.54	865.1 (0.1) ⁸⁵
Ne 1s - 3p	866.58	866.96	867.30	867.40	867.29 ⁸⁶
MSE	0.15	-0.09	0.12	0.11	
MAE	0.43	0.15	0.14	0.34	
RMSE	0.52	0.19	0.18	0.41	
MAX	1.01	0.41	0.37	0.87	

challenging case for $\Delta\text{SCF}(\text{SCAN})$ is Be, over estimated by 0.51 eV. Somewhat surprisingly $\Delta\text{SCF}(\text{HF})$ predicts the Be experimental ionization perfectly.

The performance of $\Delta\text{SCF}(\text{HF})$ against the much more sophisticated FC-CVS-EOM-CCSD-IP is once again remarkable, with the MAE and RMSE of the latter being 0.35 and 0.45 eV. These FC-CVS-EOM-CCSD-IP errors are roughly five times smaller than those reported by Liu et al.⁴⁷ We wonder whether a situation similar to that reported by Vidal et al.²³ in the context of EOM-CCSD-EE for core excitations is taking place, where the specifics of the CVS implementation resulted in differences of over 1 eV.

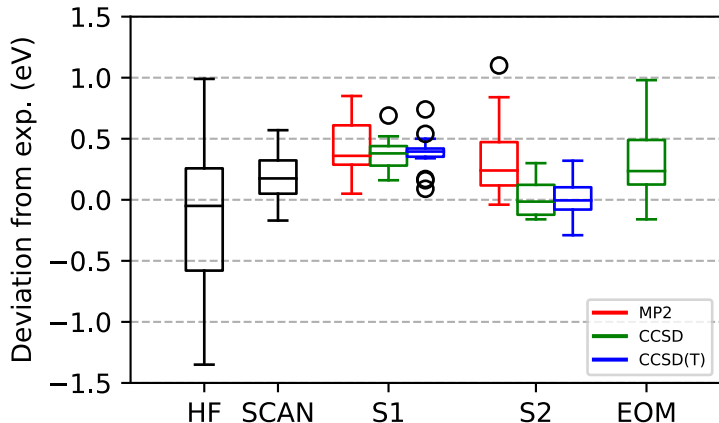


Figure 4: Statistical summary of the accuracy of calculated K-shell core ionizations relative to experimental values for the 18 ionizations shown in Table 2, as evaluated by ROHF(SCF), the correlated Δ methods (Schemes S1 and S2), and FC-CVS-EOM-CCSD-IP. For the S1 and S2 approaches, in addition to CCSD itself, the corresponding MP2 and CCSD(T) values are also shown. The specific values corresponding to these statistics are given in Table 2 and the Supplementary Information.

In contrast to excitations, the correlated Δ methods using the S1 model manage to slightly improve upon ΔHF for ionization. $\Delta\text{MP2}(\text{S1})$ increases the HF ionization energy in almost all cases, and over 1 eV in several of them: H_2CO , CH_3OH , CO , HF , F_2 , and Ne . The only case where $\Delta\text{MP2}(\text{S1})$ decreases the ionization predicted by ΔHF is CO , which is also the second most challenging case for ΔHF , right after F_2 . The problematic Be is overestimated by 0.81 eV by $\Delta\text{MP2}(\text{S1})$. Once again, $\Delta\text{CCSD}(\text{S1})$ alleviates the worst cases in $\Delta\text{MP2}(\text{S1})$. CO is anomalous in that this is the only case where $\Delta\text{CCSD}(\text{S1})$ significantly

worsens the $\Delta\text{MP2}(\text{S1})$ result, and also the only one where the (T) seems to significantly improve the result, correcting the $\Delta\text{CCSD}(\text{S1})$ result by 0.17 eV. Overall, the S1 methods result in MAEs and RMSEs of 0.42, 0.37, 0.38 eV and 0.49, 0.39, 0.41 eV for MP2, CCSD, and CCSD(T). As Lubijic²⁹ noted in their study, $\Delta\text{MP2}(\text{S1})$ seldom warrants the additional cost over ΔSCF and neither extending to CCSD or CCSD(T) seems to improve the results to an extent that justifies their cost. As for excitations, a consistent overestimation of the core ionization energies, as evidenced by the MSEs being equal to the MAEs for all the S1 correlated methods, hints at the configurations neglected by the S1 scheme being important.

Indeed, the improvement in calculated core ionization energies provided by the correlated methods under model S2, relative to S1, is even more dramatic than it is for the excitations. In contrast with $\Delta\text{MP2}(\text{S1})$, $\Delta\text{MP2}(\text{S2})$ manages to somewhat improve the statistics from ΔHF , bringing down the MAE and RMSE to 0.33 and 0.44 eV. S2 improves the S1 results for MP2 in almost all cases, the only significant exception being Be, where $\Delta\text{MP2}(\text{S2})$ performs the worst: an overestimation of 1.1 eV. As with S1, $\Delta\text{CCSD}(\text{S2})$ alleviates the failures of $\Delta\text{MP2}(\text{S2})$ (significantly for Be) and brings the MAE and RMSE down to 0.12 and 0.15 eV. $\Delta\text{CCSD}(\text{T})$ slightly worsens the statistics by bringing the MAE and RMSE to 0.13 and 0.17 eV. The RMSE for $\Delta\text{CCSD}(\text{S2})$ is more than 2.5 times smaller than for FC-CVS-EOM-CCSD-IP.

The results presented here are comparable to those in Table 5 of Zheng *et al.*³⁰ The differences can be associated with the different basis sets used and the way we are treating the correlation associated with the core virtual. Whereas in their study, they make estimates to the correlation missing due to freezing the core orbital completely (S1) by carrying out unconstrained (S0) calculations with denominator thresholds, S2 recovers it by a well-defined protocol.

Table 2: Core ionizations

Transition	Δ SCF(HF)	Δ SCF(SCAN)	Δ CCSD(S2)	EOM-CCSD	Experiment
Be 1s - ion	123.35	123.92	123.65	123.49	123.35 ⁸⁰
C ₂ H ₄ 1s - ion	290.71	290.92	290.72	290.95	290.88 ⁸⁷
CH ₄ 1s - ion	290.86	290.92	290.69	290.68	290.83 ⁸⁷
C ₂ H ₂ 1s - ion	291.39	291.47	291.21	291.26	291.14 ⁸⁷
CH ₃ OH 1s - ion	292.63	292.63	292.44	292.52	292.3 (0.2) ⁸⁸
HCN 1s - ion	293.76	293.68	293.43	293.34	293.50 ⁸⁷
H ₂ CO 1s - ion	294.91	294.75	294.50	294.70	294.35 ⁷⁸
CO 1s - ion	297.23	296.58	296.47	296.43	296.24 ⁸⁷
NH ₃ 1s - ion	405.48	405.70	405.51	405.77	405.52 ⁸⁷
HCN 1s - ion	406.74	406.96	406.78	407.10	406.8 ⁸⁷
N ₂ 1s - ion	410.21	410.15	409.99	409.89	409.9 ⁸⁷
CH ₃ OH 1s - ion	538.43	539.08	538.90	539.64	539.06 (0.2) ⁸⁸
H ₂ CO 1s - ion	538.51	539.47	539.29	540.28	539.30 ⁷⁸
H ₂ O 1s - ion	539.49	539.96	539.82	540.29	539.92 ⁸⁷
CO 1s - ion	541.79	542.65	542.43	543.10	542.57 ⁸⁷
HF 1s - ion	693.62	694.30	694.25	694.80	694.0 ⁸⁷
F ₂ 1s - ion	695.36	696.54	696.58	697.58	696.71 ⁸⁷
Ne 1s - ion	869.54	870.21	870.31	870.49	870.33 ⁸⁶
MSE	-0.15	0.18	0.02	0.31	
MAE	0.45	0.21	0.13	0.35	
RMSE	0.58	0.25	0.17	0.45	
MAX	1.35	0.57	0.33	0.98	

Conclusions

We have studied the use of core-hole orbital-optimized references in SR correlated methods to describe core excited and core ionized states of 18 small closed-shell organic molecules, and compared them against two of the most successful approaches so far: ROKS(SCF) and EOM-CC. The use of three different schemes (S1, S2, S3) to address the convergence problems of the CC equations, and the spin contamination of the excited states, were employed. S1 excludes all amplitudes involving the half-occupied core orbital associated with the excitation or ionization. S2 allows for the ones that retain a core occupancy of 1. S3, exclusively for CCSD core excitations, fixes the T_2 amplitude associated with the spin complement of a spin symmetry-broken core-excited reference to ± 1.0 , thereby ensuring the proper reference CSF is present in the cluster expansion. As evidenced by the energetic difference between the singlet and the triplet core excited states, addressing the spin contamination associated with using a symmetry broken reference is essential for quantitative studies using the correlated Δ methods unless Rydberg states are being targeted.

To compare with experimental core excitations and ionizations requires careful attention to basis set convergence, which we have addressed by using the aug-cc-pCVXZ basis set for heavy atoms ($n = T, Q$, with extrapolation), and aug-cc-pVDZ for hydrogen. With this protocol, Δ CCSD(S3) performs the best among the correlated Δ methods for core excitations, reaching an MAE and RMSE of 0.15 and 0.19 eV for CCSD. These statistics are on par with the most successful orbital-optimized DFT approach, ROKS(SCAN). Δ CCSD(S2) follows closely behind, with an MAE and RMSE of 0.18 and 0.22 eV. As such, Δ CCSD with either S2 or S3 roughly halves the errors of FC-CVS-EOM-CCSD-EE. A similar situation takes place for ionizations, where S2 in conjunction with CCSD performs the best, by achieving a MAE and RMSE of 0.13 and 0.17 eV, respectively. Δ CCSD(S2) reduces the FC-CVS-EOM-CCSD-IP error by more than a factor of 2.5, and outperforms Δ SCF(SCAN), which has an MAE and RMSE of 0.21 and 0.25 eV.

The use of a CVS scheme like S1 for the correlated Δ methods is discouraged, if quanti-

tative agreement is sought after. Furthermore, as has previously been concluded by others,²⁹ we cannot recommend the use of Δ MP2 for the prediction of core excitations or ionizations. In the future, it may be interesting to explore whether regularization^{32,89} can address the limitations of Δ MP2. Finally, we note that the use of the perturbative (T) triples correction with the best scheme that allows for it, S2, does not seem to offer a significant improvement over CCSD. Perhaps this is because the effect of triples is small (based on the excellent results obtained with Δ CCSD(S2) and Δ CCSD(S3)) or perhaps a full triples treatment is needed to obtain further significant improvement.

Considering the tractability of our Δ CCSD schemes S2 and S3, as well as the challenge of approaching the complete basis set limit, it is indeed encouraging that our results with Δ CCSD(S2) and Δ CCSD(S3) are attaining errors that approach the given experimental uncertainties (typically on the order of 0.1 eV). Apart from the BSL, the use of atomic relativistic corrections, and the truncation of correlation at the CCSD and CCSD(T) level, there are two (presumably small) additional energetic effects that we have neglected. First is the coupling of the resonance with the Auger continuum⁹⁰ and second is vibronic effects.⁹¹⁻⁹⁴

Acknowledgement

The authors thank Diptarka Hait for fruitful discussions. J. L. thanks David Reichman for support. This work was supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy, under Contract No. DE-AC02-05CH11231.

Supporting Information Available

All the Supporting Information is provided as .xlsx spreadsheets. For both excitations and ionizations via the three Δ CC schemes, the following data is provided: SCF energies, correlation energies, SCF and CCSD $\langle S^2 \rangle$ values. For the ROKS(SCAN) and FC-CVS-EOM-EE-CCSD and FC-CVS-EOM-IP-CCSD calculations, the excitation energies are provided.

References

- (1) Yu, L. H.; Shaftan, T. Towards coherent X-ray free-electron lasers. *Nature Photonics* 2019 13:8 **2019**, 13, 513–515.
- (2) Halavanau, A.; Benediktovitch, A.; Lutman, A. A.; DePonte, D.; Cocco, D.; Rohringer, N.; Bergmann, U.; Pellegrini, C. Population inversion X-ray laser oscillator. *Proceedings of the National Academy of Sciences of the United States of America* **2020**, 117, 15511–15516.
- (3) Kumar Maroju, P. et al. Attosecond pulse shaping using a seeded free-electron laser. *386 — Nature — 2005*, 578.
- (4) Duris, J. et al. Tunable isolated attosecond X-ray pulses with gigawatt peak power from a free-electron laser. *Nature Photonics* 2019 14:1 **2019**, 14, 30–36.
- (5) Chang, K. F.; Reduzzi, M.; Wang, H.; Poullain, S. M.; Kobayashi, Y.; Barreau, L.; Prendergast, D.; Neumark, D. M.; Leone, S. R. Revealing electronic state-switching at conical intersections in alkyl iodides by ultrafast XUV transient absorption spectroscopy. *Nature Communications* 2020 11:1 **2020**, 11, 1–7.
- (6) Loh, Z. H. et al. Observation of the fastest chemical processes in the radiolysis of water. *Science* **2020**, 367, 179–182.
- (7) Lin, Y. C.; Fidler, A. P.; Sandhu, A.; Lucchese, R. R.; McCurdy, C. W.; Leone, S. R.; Neumark, D. M. Coupled nuclear–electronic decay dynamics of O₂ inner valence excited states revealed by attosecond XUV wave-mixing spectroscopy. *Faraday Discussions* **2021**, 228, 537–554.
- (8) Sette, F.; Krisch, M. H.; Masciovecchio, C.; Ruocco, G.; Monaco, G. Dynamics of glasses and glass-forming liquids studied by inelastic X-ray scattering. *Science* **1998**, 280, 1550–1555.

- (9) Attar, A. R.; Chang, H.-T.; Britz, A.; Zhang, X.; Lin, M.-F.; Krishnamoorthy, A.; Linker, T.; Fritz, D.; Neumark, D. M.; Kalia, R. K.; Nakano, A.; Ajayan, P.; Vashishta, P.; Bergmann, U.; Leone, S. R. Simultaneous Observation of Carrier-Specific Redistribution and Coherent Lattice Dynamics in 2H-MoTe₂ with Femtosecond Core-Level Spectroscopy. **2020**,
- (10) Mazza, T. et al. Mapping Resonance Structures in Transient Core-Ionized Atoms. *Physical Review X* **2020**, *10*, 041056.
- (11) Haynes, D. C. et al. Clocking Auger electrons. *Nature Physics* *2021* *17:4* **2021**, *17*, 512–518.
- (12) Popmintchev, T. et al. Bright coherent ultrahigh harmonics in the keV x-ray regime from mid-infrared femtosecond lasers. *Science* **2012**, *336*, 1287–1291.
- (13) Zimmermann, P.; Peredkov, S.; Abdala, P. M.; DeBeer, S.; Tromp, M.; Müller, C.; van Bokhoven, J. A. Modern X-ray spectroscopy: XAS and XES in the laboratory. *Coordination Chemistry Reviews* **2020**, *423*, 213466.
- (14) Barreau, L.; Ross, A. D.; Garg, S.; Kraus, P. M.; Neumark, D. M.; Leone, S. R. Efficient table-top dual-wavelength beamline for ultrafast transient absorption spectroscopy in the soft X-ray region. *Scientific Reports* **2020**, *10*, 1–9.
- (15) Scutelnic, V.; Tsuru, S.; Pápai, M.; Yang, Z.; Epshtein, M.; Xue, T.; Haugen, E.; Kobayashi, Y.; Krylov, A. I.; Møller, K. B.; Coriani, S.; Leone, S. R. X-ray transient absorption reveals the 1Au ($n\pi^*$) state of pyrazine in electronic relaxation. *Nature Communications* **2021**, *12*, 6–13.
- (16) Epshtein, M.; Scutelnic, V.; Yang, Z.; Xue, T.; Vidal, M. L.; Krylov, A. I.; Coriani, S.; Leone, S. R. Table-Top X-ray Spectroscopy of Benzene Radical Cation. *Journal of Physical Chemistry A* **2020**, *124*, 9524–9531.

- (17) Besley, N. A.; Asmuruf, F. A. Time-dependent density functional theory calculations of the spectroscopy of core electrons. *Physical Chemistry Chemical Physics* **2010**, *12*, 12024–12039.
- (18) Lyon, K.; Preciado-Rivas, R.; Zamora-Ledezma, C. Charge transfer in time-dependent density functional theory. *Journal of Physics: Condensed Matter* **2017**, *29*, 423001.
- (19) Maitra, N. T. Charge transfer in time-dependent density functional theory. *Journal of physics. Condensed matter : an Institute of Physics journal* **2017**, *29*.
- (20) Maitra, N. T. Double and Charge-Transfer Excitations in Time-Dependent Density Functional Theory. <https://doi.org/10.1146/annurev-physchem-082720-124933> **2021**, *8*, 17.
- (21) Coriani, S.; Christiansen, O.; Fransson, T.; Norman, P. Coupled-cluster response theory for near-edge x-ray-absorption fine structure of atoms and molecules. *PHYSICAL REVIEW A* **2012**, *85*, 22507.
- (22) 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. *The Journal of Chemical Physics* **2015**, *142*, 214104.
- (23) Vidal, M. L.; Feng, X.; Epifanovsky, E.; Krylov, A. I.; Coriani, S. New and Efficient Equation-of-Motion Coupled-Cluster Framework for Core-Excited and Core-Ionized States. *Journal of Chemical Theory and Computation* **2019**, *15*, 3117–3133.
- (24) Triguero, L.; Plashkevych, O.; Pettersson, L. G.; Ågren, H. Separate state vs. transition state Kohn-Sham calculations of X-ray photoelectron binding energies and chemical shifts. *Journal of Electron Spectroscopy and Related Phenomena* **1999**, *104*, 195–207.
- (25) Hait, D.; Head-Gordon, M. Excited State Orbital Optimization via Minimizing the Square of the Gradient: General Approach and Application to Singly and Doubly

- Excited States via Density Functional Theory. *Journal of Chemical Theory and Computation* **2020**, *16*, 1699–1710.
- (26) Besley, N. A.; Gilbert, A. T.; Gill, P. M. Self-consistent-field calculations of core excited states. *Journal of Chemical Physics* **2009**, *130*.
- (27) Duflot, D.; Flament, J. P.; Heinesch, J.; Hubin-Franskin, M. J. The K-shell spectra of tetrahydrofuran studied by electron energy loss spectroscopy and ab initio calculations. *Chemical Physics Letters* **2010**, *495*, 27–32.
- (28) Shim, J.; Klobukowski, M.; Barysz, M.; Leszczynski, J. Calibration and applications of the Δ MP2 method for calculating core electron binding energies. *Physical Chemistry Chemical Physics* **2011**, *13*, 5703–5711.
- (29) Ljubić, I. Reliability of density functional and perturbation theories for calculating core-ionization spectra of free radicals. *Journal of Chemical Theory and Computation* **2014**, *10*, 2333–2343.
- (30) Zheng, X.; Cheng, L. Performance of Delta-Coupled-Cluster Methods for Calculations of Core-Ionization Energies of First-Row Elements. *Journal of Chemical Theory and Computation* **2019**, *15*, 4945–4955.
- (31) Zheng, X.; Liu, J.; Doumy, G.; Young, L.; Cheng, L. Hetero-site Double Core Ionization Energies with Sub-electronvolt Accuracy from Delta-Coupled-Cluster Calculations. *Journal of Physical Chemistry A* **2020**, *124*, 4413–4426.
- (32) Lee, J.; Small, D. W.; Head-Gordon, M. Excited states via coupled cluster theory without equation-of-motion methods: Seeking higher roots with application to doubly excited states and double core hole states. *Journal of Chemical Physics* **2019**, *151*.
- (33) Huang, M.; Li, C.; Evangelista, F. A. Theoretical Calculation of Core-Excited States

- along Dissociative Pathways beyond Second-Order Perturbation Theory. *Journal of Chemical Theory and Computation* **2021**, acs.jctc.1c00884.
- (34) Hu, C. H.; Chong, D. P. Density functional computations for inner-shell excitation spectroscopy. *Chemical Physics Letters* **1996**, *262*, 729–732.
- (35) Triguero, L.; Pettersson, L.; Ågren, H. Calculations of near-edge x-ray-absorption spectra of gas-phase and chemisorbed molecules by means of density-functional and transition-potential theory. *Physical Review B - Condensed Matter and Materials Physics* **1998**, *58*, 8097–8110.
- (36) Michelitsch, G. S.; Reuter, K. Efficient simulation of near-edge X-ray absorption fine structure (NEXAFS) in density-functional theory: Comparison of core-level constraining approaches. *Journal of Chemical Physics* **2019**, *150*.
- (37) Brabec, J.; Bhaskaran-Nair, K.; Govind, N.; Pittner, J.; Kowalski, K. Communication: Application of state-specific multireference coupled cluster methods to core-level excitations. *Journal of Chemical Physics* **2012**, *137*, 1–5.
- (38) Sen, S.; Shee, A.; Mukherjee, D. A study of the ionisation and excitation energies of core electrons using a unitary group adapted state universal approach. *Molecular Physics* **2013**, *111*, 2625–2639.
- (39) Dutta, A. K.; Gupta, J.; Vaval, N.; Pal, S. Intermediate Hamiltonian Fock space multireference coupled cluster approach to core excitation spectra. *Journal of Chemical Theory and Computation* **2014**, *10*, 3656–3668.
- (40) Maganas, D.; Kowalska, J. K.; Nooijen, M.; Debeer, S.; Neese, F. Comparison of multireference ab initio wavefunction methodologies for X-ray absorption edges: A case study on [Fe(II/III)Cl₄]^{2-/1-} molecules. *Journal of Chemical Physics* **2019**, *150*.

- (41) Garner, S. M.; Neuscamman, E. Core excitations with excited state mean field and perturbation theory. *The Journal of Chemical Physics* **2020**, *153*, 154102.
- (42) Garner, S. M.; Neuscamman, E. A variational Monte Carlo approach for core excitations. *The Journal of Chemical Physics* **2020**, *153*, 144108.
- (43) Gilbert, A. T.; Besley, N. A.; Gill, P. M. Self-consistent field calculations of excited states using the maximum overlap method (MOM). *Journal of Physical Chemistry A* **2008**, *112*, 13164–13171.
- (44) Barca, G. M.; Gilbert, A. T.; Gill, P. M. Simple models for difficult electronic excitations. *J. Chem. Theory Comput.* **2018**, *14*, 1501–1509.
- (45) Carter-Fenk, K.; Herbert, J. M. State-Targeted Energy Projection: A Simple and Robust Approach to Orbital Relaxation of Non-Aufbau Self-Consistent Field Solutions. *Journal of Chemical Theory and Computation* **2020**, *16*, 5067–5082.
- (46) Boys, S. F. Construction of some molecular orbitals to be approximately invariant for changes from one molecule to another. *Reviews of Modern Physics* **1960**, *32*, 296–299.
- (47) Liu, J.; Matthews, D.; Coriani, S.; Cheng, L. Benchmark Calculations of K-Edge Ionization Energies for First-Row Elements Using Scalar-Relativistic Core-Valence-Separated Equation-of-Motion Coupled-Cluster Methods. *Journal of Chemical Theory and Computation* **2019**, *15*, 1642–1651.
- (48) Oosterbaan, K. J.; White, A. F.; Hait, D.; Head-Gordon, M. Generalized single excitation configuration interaction: An investigation into the impact of the inclusion of non-orthogonality on the calculation of core-excited states. *Physical Chemistry Chemical Physics* **2020**, *22*, 8182–8192.
- (49) Yamaguchi, K.; Jensen, F.; Dorigo, A.; Houk, K. A spin correction procedure for

- unrestricted Hartree-Fock and Möller-Plesset wavefunctions for singlet diradicals and polyradicals. *Chem. Phys. Lett.* **1988**, *149*, 537–542.
- (50) Kitagawa, Y.; Saito, T.; Nakanishi, Y.; Kataoka, Y.; Matsui, T.; Kawakami, T.; Okumura, M.; Yamaguchi, K. Spin contamination error in optimized geometry of singlet carbene ($1A_1$) by broken-symmetry method. *Journal of Physical Chemistry A* **2009**, *113*, 15041–15046.
- (51) Frank, I.; Hutter, J.; Marx, D.; Parrinello, M. Molecular dynamics in low-spin excited states. *J. Chem. Phys.* **1998**, *108*, 4060–4069.
- (52) Filatov, M.; Shaik, S. A spin-restricted ensemble-referenced Kohn-Sham method and its application to diradicaloid situations. *Chemical Physics Letters* **1999**, *304*, 429–437.
- (53) Kowalczyk, T.; Tsuchimochi, T.; Chen, P.-T.; Top, L.; Van Voorhis, T. Excitation energies and Stokes shifts from a restricted open-shell Kohn-Sham approach. *J. Chem. Phys.* **2013**, *138*, 164101.
- (54) Hait, D.; Head-Gordon, M. Orbital optimized density functional theory for electronic excited states. *J. Phys. Chem. Lett.* **2021**, *12*, 4517–4529.
- (55) Hait, D.; Haugen, E. A.; Yang, Z.; Oosterbaan, K. J.; Leone, S. R.; Head-Gordon, M. Accurate prediction of core-level spectra of radicals at density functional theory cost via square gradient minimization and recoupling of mixed configurations. *J. Chem. Phys.* **2020**, *153*, 134108.
- (56) Hait, D.; Head-Gordon, M. Highly Accurate Prediction of Core Spectra of Molecules at Density Functional Theory Cost: Attaining Sub-electronvolt Error from a Restricted Open-Shell Kohn-Sham Approach. *Journal of Physical Chemistry Letters* **2020**, *11*, 775–786.

- (57) Cunha, L. A.; Hait, D.; Kang, R.; Mao, Y.; Head-Gordon, M. Relativistic Orbital Optimized Density Functional Theory for Accurate Core-Level Spectroscopy. *arXiv preprint arXiv:2111.08405* **2021**,
- (58) , L. M.; Balkova, A.; Bartlett, R. J. Multiple solutions of the single-reference coupled-cluster method. *CHEMICAL PHYSICS LETTERS* **1993**, *2*, 12.
- (59) Matthews, D. A. EOM-CC methods with approximate triple excitations applied to core excitation and ionisation energies. *Molecular Physics* **2020**, *118*.
- (60) Simons, M.; Matthews, D. A. Transition-potential coupled cluster. *Journal of Chemical Physics* **2021**, *154*.
- (61) Simons, M.; Matthews, D. A. Transition-Potential Coupled Cluster II: Optimization of the Core Orbital Occupation Number. 2022.
- (62) Stanton, J. F.; Gauss, J.; Watts, J. D.; Bartlett, R. J. A direct product decomposition approach for symmetry exploitation in many-body methods. I. Energy calculations. *The Journal of Chemical Physics* **1991**, *94*, 4334–4345.
- (63) Epifanovsky, E. et al. Software for the frontiers of quantum chemistry: An overview of developments in the Q-Chem 5 package. *The Journal of Chemical Physics* **2021**, *155*, 084801.
- (64) of Standards, N. I.; Technology, *Security Requirements for Cryptographic Modules*; 2001.
- (65) Takahashi, O. Relativistic corrections for single- and double-core excitation at the K- and L-edges from Li to Kr. *Computational and Theoretical Chemistry* **2017**, *1102*, 80–86.
- (66) Ambroise, M. A.; Jensen, F. Probing Basis Set Requirements for Calculating Core

- Ionization and Core Excitation Spectroscopy by the δ Self-Consistent-Field Approach. *Journal of Chemical Theory and Computation* **2019**, *15*, 325–337.
- (67) Ambroise, M. A.; Dreuw, A.; Jensen, F. Probing Basis Set Requirements for Calculating Core Ionization and Core Excitation Spectra Using Correlated Wave Function Methods. *Journal of Chemical Theory and Computation* **2021**, *17*, 2832–2842.
- (68) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. Basis-set convergence of correlated calculations on water. *Journal of Chemical Physics* **1997**, *106*, 9639–9646.
- (69) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. Basis-set convergence in correlated calculations on Ne, N₂, and H₂O. *Chemical Physics Letters* **1998**, *286*, 243–252.
- (70) Carbone, J. P.; Cheng, L.; Myhre, R. H.; Matthews, D.; Koch, H.; Coriani, S. *Advances in Quantum Chemistry*, 1st ed.; Elsevier Inc., 2019; Vol. 79; pp 241–261.
- (71) Woon, D. E.; Dunning Jr, T. H. Gaussian basis sets for use in correlated molecular calculations. V. Core-valence basis sets for boron through neon. *J. Chem. Phys.* **1995**, *103*, 4572–4585.
- (72) Peterson, K. A.; Dunning Jr, T. H. Accurate correlation consistent basis sets for molecular core–valence correlation effects: The second row atoms Al–Ar, and the first row atoms B–Ne revisited. *J. Chem. Phys.* **2002**, *117*, 10548–10560.
- (73) Oliphant, N.; Adamowicz, L. Multireference coupled-cluster method using a single-reference formalism. **1991**, *1229*.
- (74) Kucharski, S. A.; Bartlett, R. J. Hilbert space multireference coupled-cluster methods. I. the single and double excitation model. *Journal of Chemical Physics* **1991**, *95*, 8227–8238.

- (75) Balková, A.; Bartlett, R. J. Coupled-cluster method for open-shell singlet states. *Chemical Physics Letters* **1992**, *193*, 364–372.
- (76) Sun, J.; Ruzsinszky, A.; Perdew, J. P. Strongly constrained and appropriately normed semilocal density functional. *Phys. Rev. Lett.* **2015**, *115*, 036402.
- (77) Shaw, D. A.; King, G. C.; Read, F. H.; Cvejanovic, D. The observation of electric-dipole-forbidden inner-shell transitions in N₂ and Ar by the electron energy-loss technique. *Journal of Physics B: Atomic and Molecular Physics* **1982**, *15*, 1785–1793.
- (78) Remmers, G.; Domke, M.; Puschmann, A.; Mandel, T.; Xue, C.; Kaindl, G.; Hudson, E.; Shirley, D. A. High-resolution K-shell photoabsorption in formaldehyde. *Physical Review A* **1992**, *46*, 3935–3944.
- (79) Sodhi, R. N.; Brion, C. E. Reference energies for inner shell electron energy-loss spectroscopy. *Journal of Electron Spectroscopy and Related Phenomena* **1984**, *34*, 363–372.
- (80) Kramida, A.; Martin, W. C. A Compilation of Energy Levels and Wavelengths for the Spectrum of Neutral Beryllium (Be I). *Journal of Physical and Chemical Reference Data* **2009**, *26*, 1185.
- (81) Hitchcock, A. P.; Brion, C. E. Carbon K-shell excitation of C₂H₂, C₂H₄, C₂H₆ and C₆H₆ by 2.5 keV electron impact. *Journal of Electron Spectroscopy and Related Phenomena* **1977**, *10*, 317–330.
- (82) Hitchcock, A. P.; Brion, C. E. Inner shell electron energy loss studies of HCN and C₂N₂. *Chemical Physics* **1979**, *37*, 319–331.
- (83) Prince, K. C.; Richter, R.; De Simone, M.; Alagia, M.; Coreno, M. Near edge x-ray absorption spectra of some small polyatomic molecules. *Journal of Physical Chemistry A* **2003**, *107*, 1955–1963.

- (84) Schirmer, J.; Trofimov, A. B.; Randall, K. J.; 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 photoabsorption spectroscopy. *Physical Review A* **1993**, *47*, 1136–1147.
- (85) Hitchcock, A. P.; Brion, C. E. K-shell excitation of HF and F₂ studied by electron energy-loss spectroscopy. *Journal of Physics B: Atomic and Molecular Physics* **1981**, *14*, 4399–4413.
- (86) Müller, A.; Bernhardt, D.; Borovik, A.; Buhr, T.; Hellhund, J.; Holste, K.; Kilcoyne, A. L. D.; Klumpp, S.; Martins, M.; Ricz, S.; Seltmann, J.; Viefhaus, J.; Schippers, S. Photoionization of Ne Atoms and Ne + Ions Near the K Edge: Precision Spectroscopy and Absolute Cross-sections. *The Astrophysical Journal* **2017**, *836*, 166.
- (87) Jolly, W. L.; Bomben, K. D.; Eyermann, C. J. Core-electron binding energies for gaseous atoms and molecules. *Atomic Data and Nuclear Data Tables* **1984**, *31*, 433–493.
- (88) Hempelmann, A.; Piancastelli, M. N.; Heiser, F.; Gessner, O.; Rüdell, A.; Becker, U. Resonant photofragmentation of methanol at the carbon and oxygen K-edge by high-resolution ion-yield spectroscopy. *Journal of Physics B: Atomic, Molecular and Optical Physics* **1999**, *32*, 2677–2689.
- (89) Shee, J.; Loipersberger, M.; Rettig, A.; Lee, J.; Head-Gordon, M. Regularized second-order Møller–Plesset theory: A more accurate alternative to conventional MP2 for non-covalent interactions and transition metal thermochemistry for the same computational cost. *J. Phys. Chem. Lett.* **2021**, *12*, 12084–12097.
- (90) Carravetta, V.; Ågren, H.; Cesar, A. Is the core photoelectron band of neon asymmetric or not? *Chemical Physics Letters* **1991**, *180*, 358–364.
- (91) Coreno, M.; De Simone, M.; Prince, K. C.; Richter, R.; Vondráček, M.; Avaldi, L.;

- Camilloni, R. Vibrationally resolved oxygen $K \rightarrow \Pi^*$ spectra of O₂ and CO. *Chemical Physics Letters* **1999**, *306*, 269–274.
- (92) Prince, K. C.; Avaldi, L.; Coreno, M.; Camilloni, R.; De Simone, M. Vibrational structure of core to Rydberg state excitations of carbon dioxide and dinitrogen oxide. *Journal of Physics B: Atomic, Molecular and Optical Physics* **1999**, *32*, 2551–2567.
- (93) De Simone, M.; Coreno, M.; Alagia, M.; Richter, R.; Prince, K. C. Inner shell excitation spectroscopy of the tetrahedral molecules CX₄ (X = H, F, Cl). *Journal of Physics B: Atomic, Molecular and Optical Physics* **2002**, *35*, 61–75.
- (94) Duflot, D.; Flament, J. P.; Giuliani, A.; Heinesch, J.; Hubin-Franskin, M. J. Core shell excitation of furan at the O1s and C1s edges: An experimental and ab initio study. *The Journal of Chemical Physics* **2003**, *119*, 8946.

TOC Graphic

