The Journal of Chemical Physics

Coupled cluster study of the x-ray absorption spectra of formaldehyde derivatives at the oxygen, carbon, and fluorine K-edges

Cite as: J. Chem. Phys. **151**, 064107 (2019); https://doi.org/10.1063/1.5097650
Submitted: 27 March 2019 . Accepted: 15 July 2019 . Published Online: 13 August 2019

Federica Frati, Frank de Groot, Javier Cerezo, Fabrizio Santoro, Lan Cheng, Rasmus Faber, and Sonia Coriani

The Journal of Chemical Physics ARTICLE scitation.org/journal/jcp

Coupled cluster study of the x-ray absorption spectra of formaldehyde derivatives at the oxygen, carbon, and fluorine K-edges

Cite as: J. Chem. Phys. 151, 064107 (2019); doi: 10.1063/1.5097650
Submitted: 27 March 2019 • Accepted: 15 July 2019 •
Published Online: 13 August 2019

Federica Frati,[1,a] Frank de Groot,[1] Javier Cerezo,[2] Fabrizio Santoro,[3] Lan Cheng,[4] Rasmus Faber,[5] and Sonia Coriani[5,b]

AFFILIATIONS

[1]Debye Institute for Nanomaterials Science, Utrecht University, Universiteitsweg 99, 3584 GC Utrecht, The Netherlands
[2]Departamento de Química, Facultad de Ciencias, Universidad Autónoma de Madrid, 28049 Madrid, Spain
[3]Istituto di Chimica dei Composti Organometallici (ICCOM–CNR), Area della Ricerca del CNR, Via Moruzzi 1, I-56124 Pisa, Italy
[4]Department of Chemistry, Krieger School of Arts and Sciences, Johns Hopkins University, Baltimore, Maryland 21218, USA
[5]DTU Chemistry, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark

[a)]f.frati@uu.nl
[b)]soco@kemi.dtu.dk

ABSTRACT

We have investigated the performance of a core-valence separated scheme within the coupled cluster (CC) hierarchy of methods CC singles (CCS), CC singles and approximate doubles (CC2), and CC singles and doubles (CCSD) in reproducing the K-edge x-ray absorption spectra of the low-Z elements carbon, oxygen, and fluorine in formaldehyde (CH_2O), carbonyl fluoride (CF_2O), formyl fluoride (CHFO), and formic acid (CHOOH). The analysis covers the entire frequency region from the first $1s \rightarrow \pi^*$ excitation to the core-ionization limit, encompassing the region of Rydberg transitions. Moreover, a simulation of the vibronic progressions in the $1s \rightarrow \pi^*$ bands of both carbon and oxygen in formaldehyde has been performed at the core-valence separated CCSD level, and the results are critically compared with highly resolved experimental data for this molecule.

I. INTRODUCTION

X-ray absorption spectroscopy (XAS) is an important tool to determine the electronic structure of molecules and solids. XAS spectra are most often measured with synchrotron radiation facilities that are a source of tunable, high-intensity, x-rays. In XAS, the x-ray photon is absorbed by a core electron that is excited to a bound state or, at higher energies, to a free electron state. Calculations of core excitations are challenging because they require computational methods that explicitly account for the excitation of core-level electrons, including relaxation and electron correlation effects. A number of theoretical methods have been proposed in order to meet these challenges; see, e.g., Refs. 1–12 and the references therein. Among these methods, Time-Dependent Density Functional Theory (TD-DFT) is one of the most commonly used methods because of its reduced computational cost and its ability to yield multiple excited states.[9] TD-DFT is a formally rigorous extension of the DFT ground-state formalism, and it is regarded as the method of choice to treat electronic excited states within a density functional framework. A main drawback of the use of TD-DFT in computing core excited spectra is the large underestimation of excitation energies, at least for conventional functionals.[5] A newly proposed, and remarkably accurate, time-independent (TI) DFT-based approach for core excitations is the variational Orthogonality Constrained Density Functional Theory method of Evangelista and co-workers.[9,10]

In the realm of electronic structure wave-function based methods, the methods rooted in a coupled cluster (CC) ansatz for the wave function,[13–16] where the correct description of the quantum mechanical system can be approached in a systematic manner, are

generally considered among the most accurate. Implemented within response theory, coupled cluster (CC) methods give a reliable theoretical framework in which relaxation effects are taken into account by means of an accurate treatment of electron correlation in both the ground and excited states. Here, the performance of the coupled cluster hierarchy of methods CCS (coupled cluster singles), CC2 (coupled cluster singles and approximate doubles), and CCSD (coupled cluster singles and doubles) with a core-valence separation (CVS) scheme applied within the manifold of excited states (in the following labeled eCVS)[17] is evaluated on the carbon, oxygen, and fluorine 1s XAS spectra of the molecules formaldehyde (CH_2O), carbonyl fluoride (CF_2O), formyl fluoride (CHFO), and formic acid (CHOOH). The analysis covers the entire K-edge region from the first absorption peak to the ionization limit, i.e., including the Rydberg excitations. For the lowest energy core-excited electronic transition at both the C and O edges of formaldehyde, we have also performed the simulation of the associated vibrational progressions since this allows for a more direct comparison with the high-resolution experiments.[18] The vibronic calculations (of the polyatomic molecule formaldehyde) also represent a more compelling stress test for the level of theory since their quality also depends on the quality of the gradients and Hessians, and hereby constitute a more detailed investigation on the accuracy of the computational method.

A. Computational methodology

Coupled cluster methods are built upon the exponential ansatz of the wave function,[19] $|\Psi\rangle = e^{\hat{T}}|0\rangle$, where $|0\rangle$ stands for the reference wave function and $\hat{T} = \sum_\mu t_\mu \tau_\mu$ is the cluster operator, which is a linear combination of the excitation operators τ_μ, each weighted by the corresponding CC amplitude, t_μ. The ground-state energy and CC amplitudes can be obtained by projection of the time-independent Schrödinger equation for the non-Hermitian similarity transformed Hamiltonian, $\hat{H}^T = e^{-\hat{T}}\hat{H}e^{\hat{T}}$, onto the reference state $|0\rangle$ and the manifold of excited states $|\mu\rangle$,

$$\langle 0|\hat{H}^T|0\rangle = E\,, \qquad \langle \mu|\hat{H}^T|0\rangle = 0. \tag{1}$$

Within CC response theory,[15,20] the CC excitation energies can be computed by determining the eigenvalues of the CC Jacobian matrix $\mathbf{A}$, whose elements are $A_{\mu\nu} = \langle \mu|\left[\hat{H}^T, \tau_\nu\right]|0\rangle$. Since the CC Jacobian is nonsymmetric, the left and right eigenvectors are not adjoint to each other, so two eigenvalue equations are solved,

$$\mathbf{A}R_j = \omega_j R_j\,, \qquad L_j\mathbf{A} = \omega_j L_j \tag{2}$$

with the biorthogonality condition $L_k R_j = \delta_{jk}$. In order to solve the above large scale nonsymmetric eigenvalue equations, variants of the Davidson[21] or Lanczos[6,7] algorithms can be adopted. In the present study, the Davidson[21] algorithm is used. Transition strengths (for electric dipole moment component α) can be determined as

$$S^{\alpha\alpha}_{0\to j} = \frac{1}{2}\left\{T^\alpha_{0j}T^\alpha_{j0} + \left(T^\alpha_{0j}T^\alpha_{j0}\right)^*\right\}, \tag{3}$$

where the left and right transition moments are given by

$$T^\alpha_{0j} = \eta^\alpha R_j + \bar{M}^j(\omega_j)\xi^\alpha, \quad T^\alpha_{j0} = L_j\xi^\alpha, \tag{4}$$

and the auxiliary Lagrangian multipliers $\bar{M}^j(\omega_j)$ are obtained from the solution of the linear equation

$$\bar{M}^j(\mathbf{A} + \omega_j\mathbf{I}) = -\mathbf{F}R_j. \tag{5}$$

See, e.g., Refs. 15, 20, and 22 for a definition of the $\mathbf{F}$ matrix and of the vectors ξ^α and η^α.

In order to obtain excitation energies and strengths of core-excited states, it is convenient to apply a core-valence separated scheme[12,17,23] to decouple the energetically high-lying core excited states from valence continuum states. In practice, this also avoids solving for the exceedingly large number of lower-lying valence excitations. Here, we have used the projection scheme proposed in Ref. 17 that enables the calculations of energies and analytic gradients with only a simple modification of a well-developed linear response or equation of motion CCSD energy and gradient code. A projector $\mathcal{P}^v_I$ is applied at each iteration in the iterative solvers to remove all vector elements not referencing at least one core orbital (or a set of selected core orbitals) I, e.g., in the Davidson case,

$$\mathcal{P}^v_I(\mathbf{A}\mathcal{P}^v_I R_j) = \omega_j\mathcal{P}^v_I R_j, \tag{6}$$

and similarly for the left eigenvectors. Further details can be found in Ref. 17. An alternative CVS scheme has been proposed in Ref. 12. We should mention that the equations that determine the Lagrange multipliers for the cluster amplitudes correlate all electrons, to be consistent with the all-electron CC treatment for the ground state. This proved to be particularly important during the determination of the multipliers for the core excited state gradient, to be able to reproduce the results of a finite difference determination of the gradient.

In the Born-Oppenheimer approximation, vibronic spectra arise from the contribution of the transitions from all populated vibrational states in the initial electronic state, $|\boldsymbol{v}_1\rangle$, to those in the final electronic state, $|\boldsymbol{v}_2\rangle$. The spectrum as a function of the radiation frequency, ω, takes the following form:[24]

$$S(\omega) = C\omega \sum_{\boldsymbol{v}_1,\boldsymbol{v}_2} \rho_{\boldsymbol{v}_1}(T)\left(|\langle \boldsymbol{v}_1|\boldsymbol{\mu}_{12}|\boldsymbol{v}_2\rangle|^2\right) \times \delta\left(\omega_{\boldsymbol{v}_2} - \omega_{\boldsymbol{v}_1} + \left(\hbar^{-1}\Delta E - \omega\right)\right), \tag{7}$$

where $\rho_{\boldsymbol{v}_1}$ are the Boltzmann weights associated with each vibrational level in the initial electronic state, $\boldsymbol{\mu}_{12}$ is the electronic transition dipole moment connecting both electronic states, $\omega_{\boldsymbol{v}_1}$ and $\omega_{\boldsymbol{v}_2}$ are the frequencies of each vibrational state on either the initial or final electronic states, and ΔE is the energy difference between the minima of each state (adiabatic energy). C is a constant whose value is ~703.3 when all quantities are in atomic units in order to provide the spectrum as a molar extinction coefficient ($S(\omega) \equiv \varepsilon(\omega)$) in the usual experimental units: $M^{-1}\,cm^{-1}$.

The above Time-Independent (TI) expression implies a sum-over-states that can become unfeasible for large systems and/or high temperatures.[24] In these cases, the spectrum can be obtained more efficiently from an alternative, Time-Dependent (TD), formulation derived by taking the Fourier transform of the Dirac delta function,[25] which leads to

$$S(\omega) = \frac{C\omega}{2\pi Z_{\boldsymbol{v}_1}}\int \chi(t,T)e^{-it(\Delta E/\hbar-\omega)}dt \tag{8}$$

where Z_{v_1} is the vibrational partition function of the initial states, $|v_1\rangle$, and $\chi(t, T)$ is the correlation function, given by

$$\chi(t,T) = \mathrm{Tr}\left[\boldsymbol{\mu}_{12} e^{-itH_2/\hbar} \boldsymbol{\mu}_{12} e^{-(\beta - it)H_1/\hbar}\right], \quad (9)$$

where Tr refers to the trace operation, $\beta = (K_B T)^{-1}$, K_B is the Boltzmann constant, T is the absolute temperature, and H_1 and H_2 are the initial-state and final-state Hamiltonians, respectively.

1. Computational details

We have considered four different systems: formaldehyde CH_2O, formic acid CHOOH, carbonyl fluoride CF_2O, and formyl fluoride CHFO. Experimental geometries from Refs. 26–29 have been used for all molecules. The hierarchy CCS (coupled-cluster singles), CC2 (coupled cluster singles and approximate doubles), and CCSD (coupled cluster singles and doubles) with the eCVS scheme,[17] available in the Dalton code,[30] have been used to compute both core excitation energies, intensities, and ionization potentials. Different correlation consistent basis sets,[31] further enriched with Rydberg-type basis functions,[32] were used. For formaldehyde, two sets of Rydberg functions were tested, namely, a $3s3p3d$ set (with quantum number n = 3 and 4) and a $7s7p$ set (with n = 2–5), both obtained according to Kaufmann's prescription.[32] The notation $(3s3p3d)_{n=3-4}$ $(7s7p)_{n=2-5}$ will be used in the following when referring to such sets. In all the plots shown in the next sessions, a rigid shift has been applied in order to align each simulated spectrum with the experimental one. The shift was determined from the energy difference Δ between the first computed and the first experimental peak.

For the vibrational structure, optimized structures for both ground and core-excited states were obtained at the (equation of motion, EOM-)CCSD/aug-cc-pCVTZ level using the same eCVS scheme, implemented in CFOUR.[33] The ground and excited state Hessians were also obtained, using a fully analytical approach for the ground state[34–37] and by using numerical differentiation of analytical gradient[38,39] for the excited states.

Vibronic spectra have been obtained adopting the harmonic approximation. For the vibrational analysis, the potential energy surfaces (PES) of the ground state and the core excited states have been modeled with a quadratic expansion around their minima (i.e., with the so-called adiabatic Hessian approach, AH[40]). Moreover, the absorption spectra have been computed in the Franck-Condon approximation (i.e., the electronic transition dipole moment $\boldsymbol{\mu}_{12}$ was assumed independent of nuclear coordinates), adopting both the time-independent[24] and the time-dependent[25] methods implemented in version 3.0 of the FCclasses code.[41]

II. RESULTS AND DISCUSSION

A. Formaldehyde

In order to determine the basis set requirements, we considered, in Fig. 1, the oxygen and carbon 1s XAS of CH_2O in comparison with the experimental spectra.[18] The main spectral features for CH_2O at both the carbon and oxygen K-edge are the 1s → π^* transition, reported in the literature at 285.59 eV and 530.82 eV, respectively. They are followed, in the higher energy part of the spectra, by a series of Rydberg transitions.[18] The family of correlation consistent basis sets[42] (cc-pVXZ) has been widely employed in accurate *ab initio* calculations because of their systematic convergence of the calculated properties toward the complete basis set limit when increasing the cardinal number X.[43] Since we want to recover a large fraction of the correlation energy, one may expect that large basis sets are needed,[44] at least of triple zeta quality. The relaxation of the core is addressed by polarization and core-correlating functions. One can also expect that the description of excited states of diffuse character will benefit from the inclusion of augmentation functions and that the addition of center-of-mass Rydberg-type functions, in particular, will allow a better description of the Rydberg region.[32]

Using the above prescriptions, the cardinal number of basis set was varied as X = D, T, Q, and in Fig. 1, the oxygen and carbon K-edge spectra of CH_2O obtained at the CCSD level with different basis sets are shown. The spectra were generated by applying a Lorentzian line shape with HWHM = 0.2 eV to the electronic stick-transitions.

The oxygen K-edge spectrum in the double ζ basis shows a positive deviation, Δ, of 3.68 eV (i.e., the core excitation energy of the reference peak is overestimated), while the triple and quadruple ζ sets have deviations of 1.72 eV and 1.59 eV, respectively. The computed spectra were therefore shifted by −Δ. The spectral shape is less affected than the absolute energy by the basis set so that even a small basis set can reasonably reproduce the experimental features. In Fig. 1, the energy separations between the 1s → π^* and the first Rydberg transition in the O edge are also reported and compared to the experimental reference value of 4.61 eV. The peak separations are less affected than the absolute energy by the variation in cardinal number, and they are overestimated by approximately 0.6 eV. We expect that this somewhat large separation is reduced if triple and higher excitations are taken into account, which would be of interest for future investigations.

In the right panels of Fig. 1, the same basis sets used for the O K-edge are tested for the carbon K-edge. The high-resolution experimental spectrum[18] shows a rich vibrational progression that will be discussed later on. Also in this case, the double ζ basis set shows the larger shift with respect to the reference experimental value (2.61 eV), and the overall shift is smaller than in the oxygen case. This appears to be a general trend, as revealed by inspection of the results for the 1s → π^* excitation of all molecules here considered, collected in Table I: the higher the energy to excite the core electron is, the larger the shift of the computed peak from the experimental counterpart is. In the case of CH_2O, both edges show a small difference in the absolute energy shift between the triple and the quadruple ζ basis set. As the overall spectral shape before the ionization limits (294.3 eV for the carbon and 539.3 eV for the oxygen threshold[18]) is not largely affected by the size of the basis set, one can possibly use a basis set of double-ζ quality if the results are then shifted accordingly. It should be noted, however, that this is partly due to the fact that all basis sets considered above include additional Rydberg functions, which play a fundamental role in the correct description of the weak bands in the spectra. This is clearly evident in Fig. 2, where we compare the spectra obtained at both carbon and oxygen K-edges using the core-valence triple ζ basis with and without Rydberg and augmentation functions. While the position of the 1s → π^* peak is roughly the same for all three basis sets, a correct description of the Rydberg region clearly requires the inclusion of diffuse functions (sufficient to describe the lower lying Rydberg

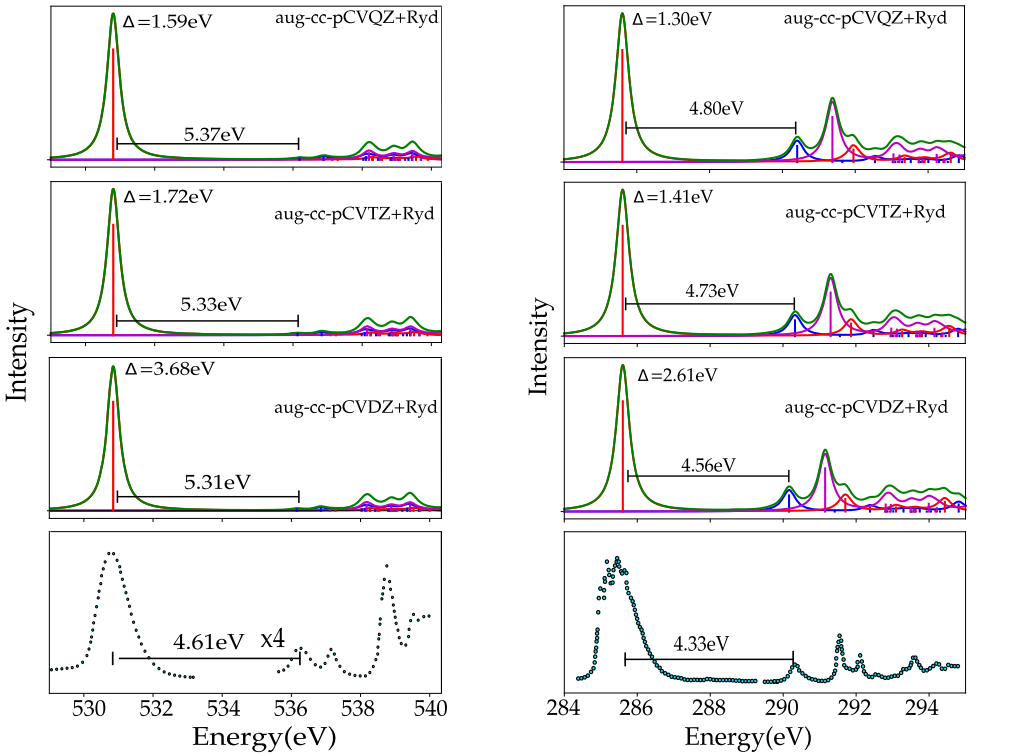


FIG. 1. Formaldehyde: eCVS-CCSD oxygen (left) and carbon (right) 1s XAS spectra in different correlation consistent basis sets supplemented with (3*s*3*p*3*d*) Rydberg functions. A parallel shift $-\Delta$ was applied to the computed spectra to align them with the first experimental peak. The total spectra are reported in green; blue lines and sticks refer to transitions of A_1 symmetry, red lines refer to those of B_1 symmetry, and magenta lines refer to those of B_2 symmetry, for the molecule placed on the *yz* plane and the C_2 axis along *z*. A Lorentzian broadening of HWHM = 0.2 eV has been applied. The experimental spectra are taken from Ref. 18. The spectral profiles shown have been generated by broadening the raw spectral data reported in Tables I and II of the supplementary material.

states) and of Rydberg-type functions (for the higher lying Rydberg states).

In Fig. 3, we show the spectra obtained for CH_2O using triple ζ basis sets with and without core polarization and/or augmentation functions, combined with two different choices of Rydberg-type basis sets. At both edges, the smallest systematic shift is found for the aug-cc-pVTZ basis set, whereas in the core-valence bases with and without augmenting functions the shifts are roughly the same. The two selected Rydberg-type basis sets yield similar spectral profiles at frequencies above the $1s \rightarrow \pi^*$ peak, though with some differences in the intensities and in the separation from the $1s \rightarrow \pi^*$ band, in particular, when combined with the cc-pCVTZ basis set. Depending on the basis set adopted, the separation between the first peak in the Rydberg region and the strong $1s \rightarrow \pi^*$ band is overestimated by approximately 0.6–0.9 eV in the case of oxygen and by ca. 0.3–0.6 eV in the case of carbon.

We note in passing that since core electrons are involved, one can expect relativistic effects to play a role in the energetics of the excitation. The inclusion of the relativistic effects by means of the Douglas-Kroll-Hess scalar relativistic Hamiltonian[45,46] and by the spin-free exact two-component theory in its one-electron variant[47,48] has been shown to give a small (positive) shift in light elements,[7,10,49–51] consistent with the relativistic contraction of inner orbitals.

Having analyzed, for CH_2O, the role of the basis set at the CCSD level, we now turn our attention to the importance of the accurate description of electron correlation within the CC hierarchy and show, in Fig. 4, the spectral profiles obtained at the O and C edges of formaldehyde within the eCVS-CCS, eCVS-CC2, and eCVS-CCSD hierarchy in the aug-cc-pCVTZ basis set supplemented with the $(3s3p3d)_{n=3-4}$ Rydberg basis functions. The computed spectra are once again compared with the experimental counterparts from Ref. 18. At the CCS level of theory (which is equivalent to configuration interaction singles, CIS), the position of the XAS first band is overestimated by as much as 15 eV at the O edge and 8.7 eV at the C edge. The next band is found at ≈9.8 eV for oxygen and ≈8.07 eV for carbon, respectively, and its intensity, relative to the first band, is too large. It is thus clear that CCS does not give a satisfactory description of the energetics of the excitations, of their relative intensities and therefore of the overall spectral features. The inadequacy of the CCS method is mainly due to its inability to describe the strong orbital relaxation effects that follow the excitation of core electrons due to the complete lack of double excitations in both ground and excited state descriptions. Indeed, when orbital relaxation is explicitly taken into account, for instance, by using as basis the electron-attached states of independently optimized, core-ionized references, (nonorthogonal) CIS-based methods have been shown to yield good agreement with experiment, at least for the $1s \rightarrow \pi^*$ excitations.[52,53]

In CC2, the (ground state) doubles amplitudes are correct to first order and the singles are correct to second order. The single-single block of the CC2 Jacobian is correct to second order, whereas the double-single and single-double blocks are correct to first order since the lowest-order coupling to the singles spectrum is retained. This ensures that the CC2 single replacement dominated excitation energies are correct through second order. Double replacement dominated excitations, on the other hand, cannot be described by CC2.[54,55]

TABLE I. Experimental (E^{exp}) and calculated (E^{cal}) 1s → π^* core-excitation energies and $1s^{-1}$ ionization energies (IE^{cal}) in eV. In the experimental case, the term values T^{exp} = (IE − E^{exp}) are given. The computed oscillator strengths (f) are also reported. Basis set aug-cc-pCVTZ (+Rydberg for CH_2O).

	C 1s → π^*					O 1s → π^*					F 1s → π^*				
System	E^{exp}	T^{exp}	E^{cal}	f	IE^{cal}	E^{exp}	T^{exp}	E^{cal}	f	IE^{cal}	E^{exp}	T^{exp}	E^{cal}	f	IE^{cal}
						eCVS-CCSD									
CH_2O	285.6[a]	8.93	287.01	0.06	296.06	530.8	9	532.52	0.045	541.68	...	...	...	...	...
CHOOH	288.5[b]	7.6	289.35	0.07	297.46	532.2	6.9	533.77	0.039	541.12	...	...	...	...	...
	...	...	...	...	...	535.3	5.4	537.56	0.017	542.71	...	...	...	...	...
CF_2O	290.9[c]	8.8	292.27	0.09	301.96	532.7	8.1	534.67	0.041	542.73	689.2	6.2	692.13	0.028	697.71
CHFO	288.2[c]	8.8	289.62	0.07	298.71	532.1	8.0	533.66	0.042	542.17	687.7	7.0	690.61	0.013	696.56
						eCVS-CC2									
CH_2O	285.6	8.9	288.52	0.06	296.02	530.8	9	532.54	0.039	538.04	...	...	...	...	...
CHOOH	288.5[b]	7.6	290.79	0.078	297.41	532.2	6.9	533.41	0.032	537.09	...	...	...	...	...
	...	...	...	...	...	535.3	5.4	536.24	0.013	539.07	...	...	...	...	...
CF_2O	290.9[c]	8.8	293.74	0.092	301.45	532.7	8.1	534.60	0.035	539.02	689.2	6.2	689.75	0.021	693.04
CHFO	288.2[c]	8.8	291.12	0.078	298.78	532.1	8	533.53	0.036	538.29	687.7	7.0	687.93	0.010	691.61
						eCVS-CCS									
CH_2O	285.6	8.9	294.29	0.11	308.63	530.8	9	545.95	0.08	559.97	...	...	...	...	...
CHOOH	288.5	7.6	295.98	0.124	310.14	532.2	6.9	547.87	0.070	559.64	...	...	...	...	...
	...	...	...	...	...	535.3	5.4	553.97	0.041	561.33	...	...	...	...	...
CF_2O	290.9[c]	8.8	298.58	0.13	313.86	532.7	8.1	548.51	0.072	561.13	689.2	6.2	710.78	0.060	718.12
CHFO	288.2[c]	8.8	296.35	0.12	308.14	532.1	8.0	547.35	0.074	558.03	687.7	7.0	708.58	0.063	716.99

[a]Experimental values for Ref. 18.
[b]Experimental values for Ref. 56.
[c]Experimental values for Ref. 57.

The CC2 oxygen K-edge spectrum of formaldehyde in Fig. 4 shows a shift of around 1.7 eV, which is basically the same as in the CCSD case. It was observed earlier[6] that CC2 can sometimes give better absolute core energies than CCSD, but this can be related to some error cancellation effects between the incomplete treatment of the double excitations and the lack of triple excitations.[7] Despite the good alignment of the first excitation energy, however, the remaining peaks in the eCVS-CC2 oxygen K edge spectrum are too close,

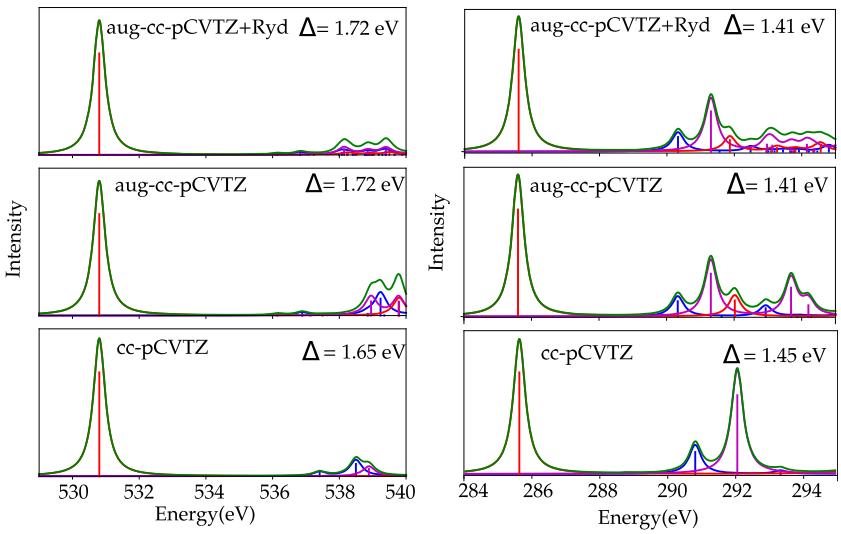


FIG. 2. Formaldehyde: eCVS-CCSD oxygen (left) and carbon (right) 1s XAS spectra using the cc-pCVTZ basis, the aug-cc-pCVTZ and the aug-cc-pCVTZ basis plus the (3*s*3*p*3*d*) Rydberg set. The parallel shifts −Δ applied to the computed spectra to align them with the first experimental peak are indicated. The total spectra are reported in green; blue lines and sticks refer to transitions of A_1 symmetry, red lines refer to those of B_1 symmetry, and magenta lines refer to those of B_2 symmetry, for the molecule placed on the *yz* plane and the C_2 axis along *z*. A Lorentzian broadening of HWHM = 0.2 eV has been applied.

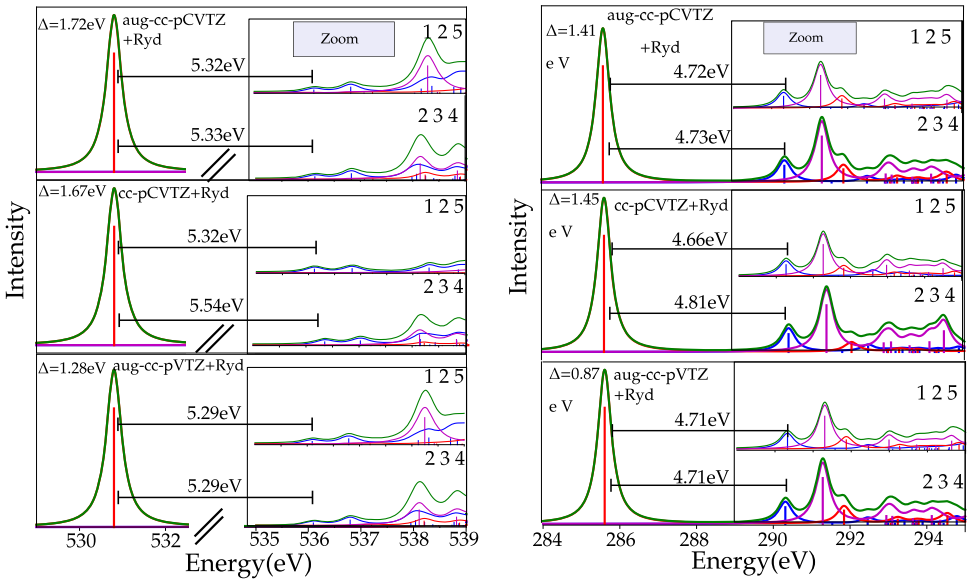


FIG. 3. Formaldehyde: eCVS-CCSD oxygen (left) and carbon (right) 1s XAS using ζ zeta basis sets with both augmentation and core correlation functions, with core correlation functions only, and with augmentation functions only. Two different Rydberg spaces are reported. A parallel shift $-\Delta$ was applied to align the computed spectra with the first experimental peak. The reported value between the first and the second peak is the difference in energy between these two features. A Lorentzian broadening of HWHM = 0.2 eV has been applied. The total spectra are in green; blue lines and sticks refer to transitions of A_1 symmetry, red lines refer to those of B_1 symmetry, and magenta lines refer to those of B_2 symmetry, for the molecule placed on the *yz* plane and the C_2 axis along *z*.

the second peak being located at only ≈2.5 eV from the first one. At the C K-edge, on the other hand, the rigid shift of the CC2 spectrum with respect to the experimental data is 2.93 eV, almost twice the value of 1.41 eV found for CCSD. Thus, at the C K-edge, the overall spectral profiles yielded by both CC2 and CCSD match rather well the experimental one, with a separation between the first and second peaks reasonably close to the (estimated) experimental value.

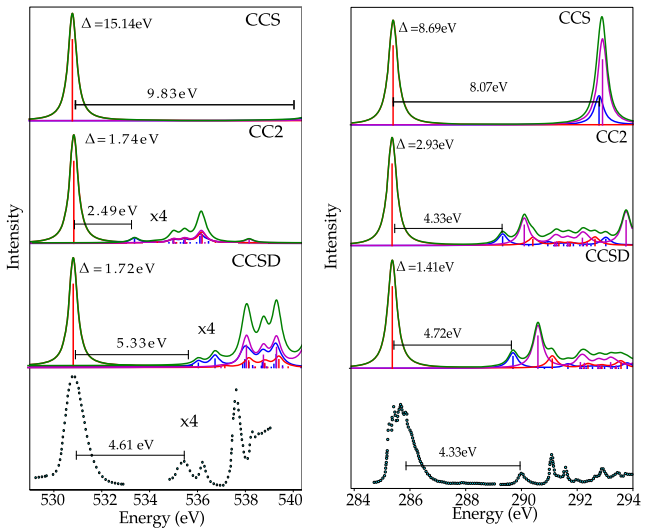


FIG. 4. Formaldehyde: eCVS-CCS, eCVS-CC2, and eCVS-CCSD oxygen (left) and carbon (right) 1s XAS spectra in the aug-cc-pCVTZ basis set supplemented with Rydberg functions, compared with the experimental results from Ref. 18. The total spectra are in green; blue lines and sticks refer to transitions of A_1 symmetry, red lines refer to those of B_1 symmetry, and magenta lines refer to those of B_2 symmetry, for the molecule placed on the *yz* plane and the C_2 axis along *z*. The spectral profiles shown have been generated by broadening the raw spectral data reported in Tables I, III, and VI of the supplementary material.

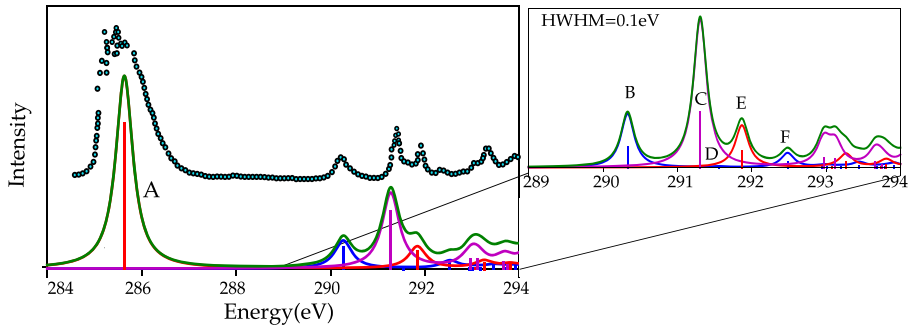
Turning our attention to the ionization energies (IEs) and (relative) term values (*T*) reported in Table I for formaldehyde, the eCVS-CCSD core ionization energy of C is ≈1.5 eV higher than the experimental value, which yields a slightly overestimated term value (9.05 eV vs the experimental value of 8.93 eV). The eCVS-CC2 core ionization energy is almost the same, but the resulting *T* is underestimated (7.5 eV). The computed oxygen K-edge ionization energy at the CCSD level yields a term value for the first excitation of 9.16 eV, vs the experimental value of 9 eV. The term value is, on the other hand, underestimated by 3.5 eV at the CC2 level mainly because of the underestimation of the core ionization energy at this level of theory. The core ionization energies and term values at both edges are strongly overestimated by eCVS-CCS.

To conclude the analysis of the purely electronic XAS spectra of formaldehyde, we report in Figs. 5 and 6 a characterization of the XAS bands obtained at the eCVS-CCSD/aug-cc-pCVTZ $+(3s3p3d)_{n=3-4}$ Rydberg level by means of the natural transition orbitals (NTOs) for the underlying core excitations. The assignments originally done by Remmers *et al.*[18] are also indicated in the figures.

For each transition, the dominant hole/particle pair is indicated; the weight of each given pair for all the transitions considered is around 0.94. The symmetry labeling indicated for each transition is the one resulting from our calculations.[18] The assignments of the peaks are mostly in line with the ones of Remmers *et al.*[18] (and Trofimov *et al.*[58]).

B. Formic acid, formyl fluoride, and carbonyl fluoride

Having discussed in greater detail both basis set and correlation effects in the case of the formaldehyde molecule in Sec. II A, we briefly consider here the other three molecular systems, for which experimental data are either more scarce or less resolved.



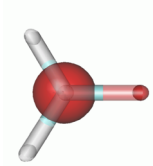
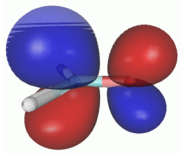
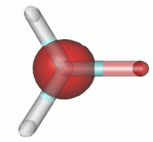
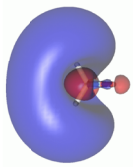
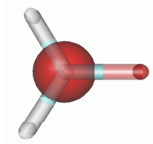
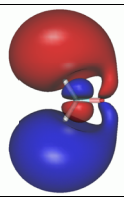
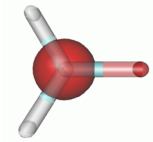
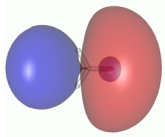
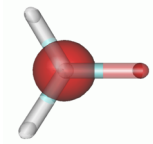
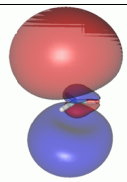
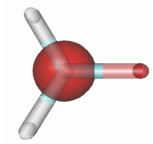
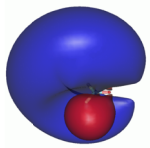
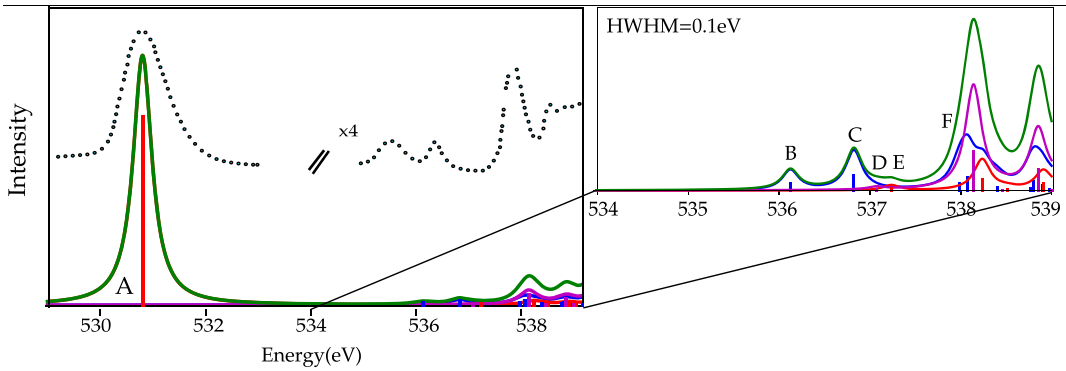
Peak	Hole	Particle	Isosurface values	Expt. assign.
A $\omega_i = 287.01$ $f \times 10^2 = 6.14$ Symmetry:B_1			Hole: 0.05 Particle: 0.05	$1s \rightarrow \pi^*$
B $\omega_i = 291.74$ $f \times 10^2 = 0.878$ Symmetry:A_1			Hole: 0.05 Particle: 0.02	$1s \rightarrow 3s\text{-}a_1$
C $\omega_i = 292.71$ $f \times 10^2 = 2.42$ Symmetry:B_2			Hole: 0.05 Particle: 0.02	$1s \rightarrow 3p\text{-}b_2$
D $\omega_i = 292.99$ $f \times 10^2 = 0.0138$ Symmetry:A_1			Hole: 0.05 Particle: 0.02	–
E $\omega_i = 293.28$ $f \times 10^2 = 0.682$ Symmetry:B_1			Hole: 0.05 Particle: 0.02	$1s \rightarrow 3p\text{-}b_1$
F $\omega_i = 293.86$ $f \times 10^2 = 0.237$ Symmetry:A_1			Hole: 0.05 Particle: 0.02	$1s \rightarrow 3d\text{-}a_1$

FIG. 5. Formaldehyde: eCVS-CCSD NTOs of the lowest C K-edge excitations. The computed spectrum has been shifted to align with the experiment.[18] The experimental assignments are from Ref. 18.



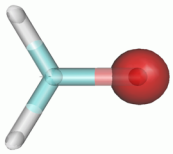
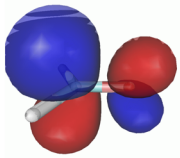
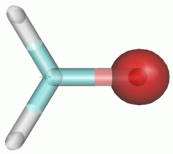
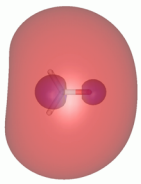
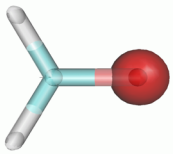
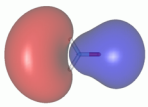
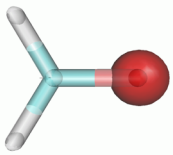
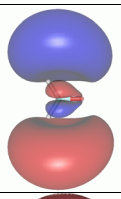
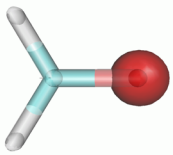
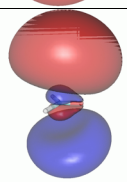
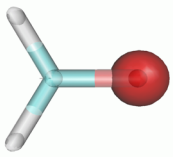
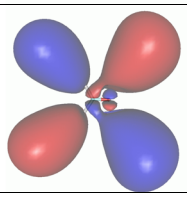
Peak	Hole	Particle	Isosurface values	Expt. assign.
A $\omega_i = 532.52$ $f \times 10^2 = 4.59$ Symmetry:B_1			Hole: 0.05 Particle: 0.05	$1s \rightarrow \pi^*$
B $\omega_i = 537.85$ $f \times 10^2 = 0.049$ Symmetry:A_1			Hole: 0.05 Particle: 0.02	$1s \rightarrow 3s\text{-}a_1$
C $\omega_i = 538.55$ $f \times 10^2 = 0.095$ Symmetry:A_1			Hole: 0.05 Particle: 0.02	$1s \rightarrow 3p\text{-}a_1$
D $\omega_i = 538.80$ $f \times 10^2 = 0.011$ Symmetry:B_2			Hole: 0.05 Particle: 0.02	–
E $\omega_i = 538.96$ $f \times 10^2 = 0.03$ Symmetry:B_1			Hole: 0.05 Particle: 0.02	–
F* $\omega_i = 539.86$ $f \times 10^2 = 0.254$ Symmetry:B_2			Hole: 0.05 Particle: 0.02	$1s \rightarrow 3d\text{-}\sigma$

FIG. 6. Formaldehyde: eCVS-CCSD NTOs of the lowest oxygen K-edge excitations. For the F band, only the NTO of the dominant B_2 component is reported. The computed spectrum has been shifted to align with the experiment.[18] The experimental assignments are from Refs. 18 and 58.

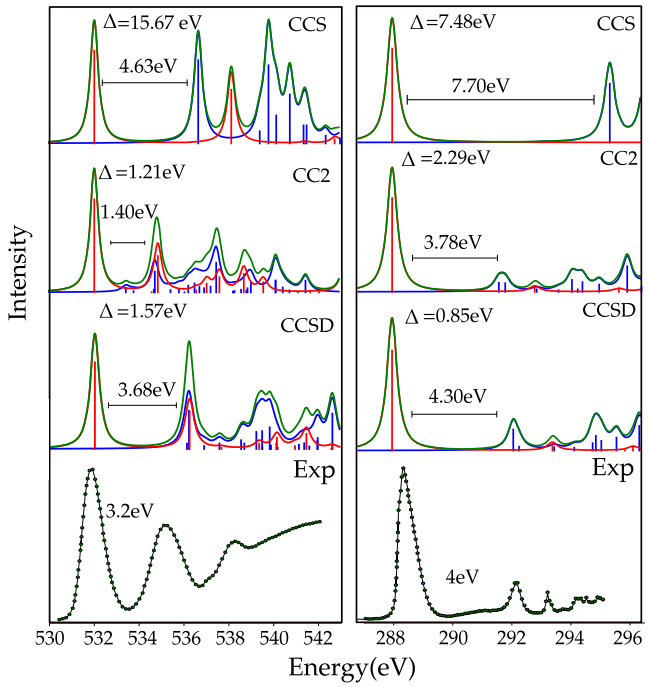


FIG. 7. Formic acid: eCVS-CCS, eCVS-CC2, and eCVS-CCSD oxygen (left) and carbon (right) 1s XAS spectra using, respectively, triple ζ basis sets with augmentation and core correlation functions, without Rydberg functions. The computed spectra were shifted by $-\Delta$ in order to align them with the first experimental peak. The experimental data are taken from Ref. 56. The total spectra are in green; blue lines and sticks refer to transitions of A′ symmetry, and red lines refer to those of A″ symmetry for the molecule with C_s symmetry. The spectral profiles shown have been generated by broadening the raw spectral data reported in Tables IX and X of the supplementary material.

As anticipated, we have collected in Table I the eCVS-CCS, eCVS-CC2, and eCVS-CCSD results for the 1s → π^* excitations, oscillator strengths, and core-ionization energies of all edges in formaldehyde, carbonyl fluoride, formyl fluoride, and formic acid obtained in the aug-cc-pCVTZ basis set (plus Rydberg for formaldehyde). Experimental absolute energies and term values T are also included. The full spectral profiles yielded by the three coupled cluster methods for CHOOH, CHFO, and CF_2O are reported in Figs. 7–9.

At the CCS level of theory, the deviation from the experiment of the absolute energy of the first excitation at the C edges varies in between 7 and 16 eV, depending on the molecules; for the oxygen edge, it is around 15–16 eV, and for fluorine, it is of the order of 21 eV. From Table I and Figs. 7 and 8, it is again clear that CCS completely fails in describing both the energetics and the overall spectral features for all systems due to its inability to describe core orbital relaxation.

In the case of formic acid, see Fig. 7, the eCVS-CC2 systematic shift Δ is smaller than the eCVS-CCSD one at the oxygen K-edge but more than twice as large at the carbon K-edge. The eCVS-CC2 spectrum at the oxygen K-edge is again more compressed than the eCVS-CCSD one, whereas at the carbon K-edge the eCVS-CC2 and eCVS-CCSD spectral profiles are very similar, at least up to 294 eV.

Figures 8 and 9 show the XAS spectra at the carbon, oxygen, and fluorine K-edges in formyl fluoride and carbonyl fluoride, respectively, obtained with the eCVS-CCS, eCVS-CC2, and eCVS-CCSD hierarchy and the aug-cc-pCVTZ basis set. As in the previous cases, eCVS-CC2 yields spectral profiles very similar to eCVS-CCSD at the carbon K-edge, whereas at the oxygen K-edge the peak separations are underestimated and the spectra appear compressed compared to eCVS-CCSD. The most striking differences between eCVS-CC2 and eCVS-CCSD are observed for the fluorine K-edges where, within the limits of the rather low experimental resolution, only eCVS-CCSD yields spectra in reasonably good agreement with the experimental ones.

Finally, as observed for formaldehyde, the core ionization energies yielded by eCVS-CCSD are always overestimated by 1.5–3.1 eV, and the eCVS-CC2 ones are underestimated by 2–4 eV, depending on the edge. The deviations are not exactly the same as the systematic shifts Δ, resulting in slightly higher (for eCVS-CCSD) or remarkably smaller (for eCVS-CC2) term values than the experimental counterparts, the difference increasing at higher energy K-edges.

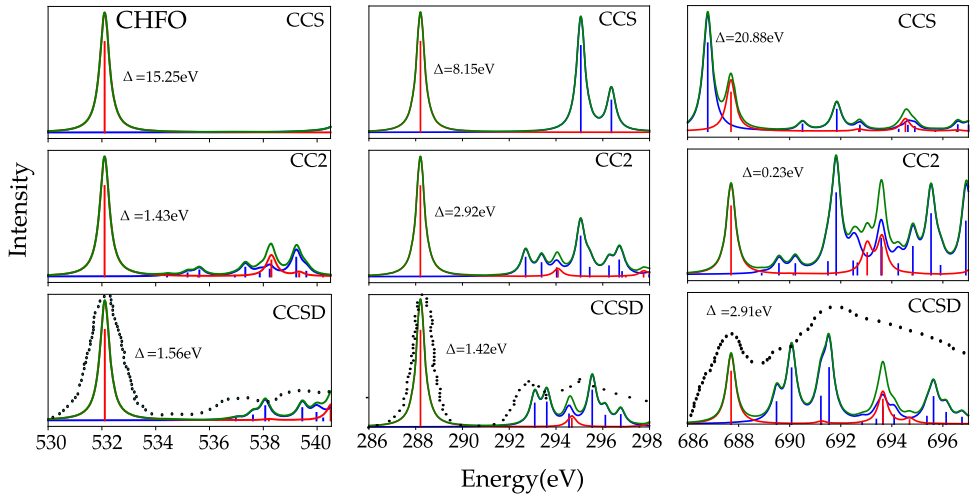


FIG. 8. CHFO: eCVS-CCSD, -CC2, and -CCS oxygen (left), carbon (middle), and fluorine (right) 1s XAS spectra using the aug-cc-pCVTZ basis set. The symbol Δ indicates the shift applied to align the computed spectra with the first experimental peak. The experimental data are taken from Ref. 59. The total spectra are in green; blue lines and sticks refer to transitions of A′ symmetry, and red lines refer to those of A″ symmetry for the molecule with C_s symmetry. The spectral profiles shown have been generated by broadening the raw spectral data reported in Tables V and VI of the supplementary material.

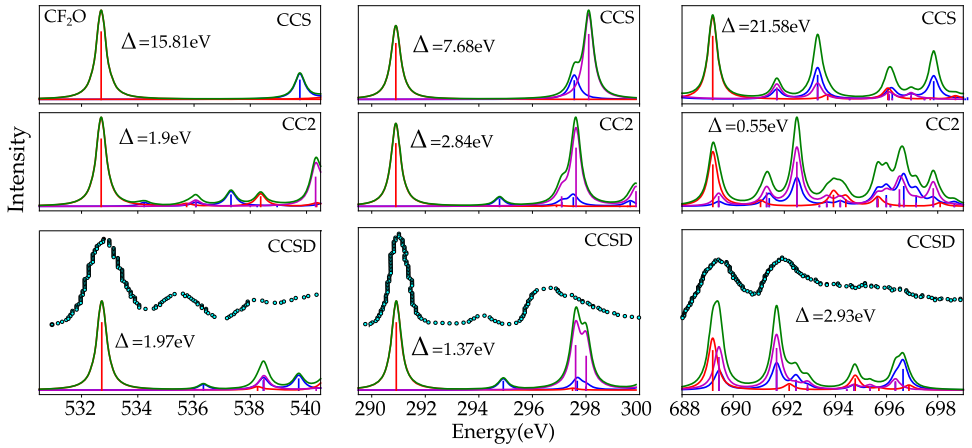


FIG. 9. CF_2O: eCVS-CCSD, eCVS-CC2, and eCVS-CCS oxygen (left), carbon (middle), and fluorine (right) 1s XAS spectra using the aug-cc-pCVTZ basis set. The symbol Δ indicates the shift applied to align the computed spectra with the first experimental peak. The experimental data are taken from Ref. 57. The total spectra are in green; blue lines and sticks refer to transitions of A_1 symmetry, red lines refer to those of B_1 symmetry, and magenta lines refer to those of B_2 symmetry, for the molecule placed on the *yz* plane and the C_2 axis along *z*. The spectral profiles shown have been generated by broadening the raw spectral data reported in Tables VII and VIII of the supplementary material.

C. Vibronic structure of the bands

As any other electronic spectroscopy, the simulation of XAS spectra directly comparable with experiments should account for the shape of the bands, which can arise from different mechanisms. The finite lifetime of the excited states induces a natural broadening, especially relevant for high energy states, as is the case in XAS spectroscopy. This contribution can be properly modeled by a Lorentzian convolution, as done in the spectra presented so far. In addition, the existence of quantum vibrational states of the chromophore leads to a number of vibronic transitions that confer a well defined structure to the bands. Finally, interaction with the environment results in the so-called inhomogeneous broadening, which can be accounted for through a Gaussian convolution.

We will focus our analysis of the spectral shape on the C $1s \rightarrow \pi^*$ and O $1s \rightarrow \pi^*$ bands in formaldehyde, which have been well characterized experimentally, including high resolution spectra.[18]

The experimental results, shown in the bottom panels of Figs. 1 and 4, indicate that such a band is characterized by a FWHM of ~1.5 eV. The shape seems structureless with this resolution, but in any case it is clearly different from the Lorentzian shape adopted to convolute the electronic sticks in the upper panels of Fig. 4. As we will show below, such broadening mainly arises from the vibrational progressions.

Let us first address the simulation of the C $1s \rightarrow \pi^*$ band. In this case, the high resolution experiments conducted by Remmes and co-workers[18] not only confirm the broadening reported in Figs. 1 and 4, but they also reveal a well resolved vibrational shape. In order to simulate the band, we computed the vibronic spectrum at the harmonic approximation adopting the AH model and computing the Hessian at ground and excited states minima with CCSD and EOM-CCSD, respectively.

The symmetry and frequencies of the most important normal modes, together with a qualitative descriptions in terms of valence internal modes, are reported in Table II. All normal modes exhibit strong frequency changes in the different states. Moreover, while ν_1 and ν_4 correspond to a unique internal coordinate, namely, the symmetric CH stretching and the out-of-plane (oop) bending, ν_2 and ν_3, arise from the combination of two modes, the CO stretch and the in-plane HCH bending, with weights markedly dependent on the electronic state. In particular, while in the ground state the CO stretching mainly contributes to ν_2, in the excited states, due to the decrease in its frequency it mixes remarkably with the HCH bending. This results in a strong Duschinsky mixing of the corresponding normal modes. In the deuterated compound, the frequency of the HCH bending is remarkably lower so that even in the excited states this mode mixes less effectively with the CO stretching.

In Fig. 10, we show the spectra simulated at 0 K for both the deuterated and nondeuterated species. As observed in this figure, our simulations do provide similar vibrational features as compared with the experimental spectra of Remmers and co-workers, although

TABLE II. Formaldehyde: most important vibrational modes and their frequencies in cm^{-1}.

Vibration	Symmetry	Description[a]	Ground state	C $1s \rightarrow \pi^*$	O $1s \rightarrow \pi^*$
ν_1	A_1	CH stretching	2965	3582	3235
ν_2	A_1	CO stretch. + HCH bend.	1823	1703	1564
ν_3	A_1	CO stretch. + HCH bend.	1550	1529	1465
ν_4	B_1	oop bending	1209	383	*i*114

[a]Note that ν_2 and ν_3 undergo strong Duschinsky mixing.

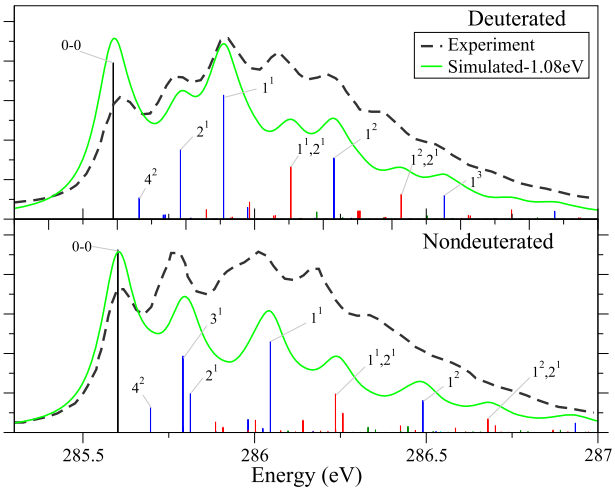


FIG. 10. Formaldehyde: simulation of the C $1s \rightarrow \pi^*$ band at 0 K for deuterated and nondeuterated species, computed on the grounds of EOM-CCSD PES, adopting a Lorentzian broadening of HWHM = 0.12 eV. Note that the simulated spectra are red-shifted by 1.08 eV so as to match the position of the experimental bands, taken from Ref. 18. Vertical lines show the vibronic transitions from the vibrational ground state in the initial electronic state. The labels correspond to the vibrational state in the final electronic states, represented as X^y, where X is the normal mode index and y is the number of quanta.

with significant differences in the relative intensity along the progression. According to our simulations, the main vibrational progressions that characterize the spectrum arise from the CH stretching (mode ν_1) and the stretching of the C=O bond (contributing to both ν_2 and ν_3). In this sense, the differences between the simulated and observed progressions can be explained by the equilibrium values of such a bond. At the (EOM-)CCSD level, the C=O equilibrium bond length at the ground and excited states is 1.202 and 1.266 Å, respectively. These values should be compared with those estimated by Remmers *et al.*,[18] which are 1.207 and 1.316 Å for ground and C $1s \rightarrow \pi^*$ states, respectively. It is thus evident that while the simulated and experimentally derived values at the ground state are in reasonable agreement, the elongation in the excited state is considerably underestimated in our EOM-CCSD calculation, which is consistent with the less pronounced progression observed in the simulated spectrum. This error can be attributed to the limitations of EOM-CCSD to properly reproduce the excited state PES. This indicates that the inclusion of triples and quadruples corrections would be of interest to further improve the results. This, however, goes beyond the present work that focuses on the assessment of the performance of the practical CCSD approach. Trofimov *et al.*[60] explored the effect on the spectral shape of a coupling, promoted by the ν_4 out-of-plane bending motion, between C $1s \rightarrow \pi^*$ and C $1s \rightarrow 3s$ states, lying ~4.6 eV above the lower one. They adopted a linear vibronic coupling (LVC) model parameterized with ADC(2) and MRCI (multireference configuration interaction) data, while the interstate coupling was fitted to reproduce a double-well profile of C $1s \rightarrow \pi^*$ along ν_4 mode, predicted with ROHF calculations.[61] They found that interstate coupling only slightly affects the spectral shapes, introducing a significant progression along the coupling mode ν_4. According to our calculations, there is a single minimum along ν_4 that preserves the C_{2v} symmetry. Notwithstanding this, we also predict a progression with even quantum numbers along the B_1 mode ν_4 due to the strong decrease in its frequency (see Fig. 10). The overall shape of the spectrum of Ref. 58 is therefore very similar even if the coupling is neglected, as we do. As far as the relative intensity of the main bands is concerned, the predictions of Refs. 58 and 60 are in better agreement with the experiment with respect to our results. This is due to the larger displacement of the equilibrium distance of the CO bond predicted at the ADC(2) level of theory (1.280 Å in the C $1s \rightarrow \pi^*$ state).[60] On the other side, the displacement of the CH stretching is pretty similar: in the excited state, it is 1.02 Å according to our calculations and 1.03 Å in Trofimov's model. Our calculations accurately reproduce the spacings of the vibronic bands, thanks to the fact that the AH model explicitly accounts for the large frequency changes undergone by CO and CH stretching (see Table II) and for the Duschinsky mixings. A calculation with the AS model[40] that, analogously to the LVC model adopted by Trofimov *et al.*, assumes that the excited-state normal modes and frequencies are equal to those of the ground state leads to clearly less satisfactory results (see Fig. 1 of the supplementary material).

The agreement with the spectrum of the deuterated species is also satisfactory. In particular, we reproduce the enhancement of the 1^1 band since, due to the larger mass of the D isotope, the dimensionless displacement increases from 1.10 (CH stretch) to 1.38 (CD stretch), and, at the same time, its frequency reduces from 3582 cm^{-1} (Table II) to 2593 cm^{-1}. On the contrary, the C=O stretching mode is practically unaltered, but, due to the decoupling with the DCD bending, the Duschinsky mixing is very limited and a single 2^1 band substitutes the 2^1, 3^1 pair in the nondeuterated species.

We now turn our attention to the O $1s \rightarrow \pi^*$ band. In this case, the excited state geometry has a lower symmetry (C_s) as compared to the ground state minimum (C_{2v}). Concretely, whilst the structure is planar in the ground state, it is partially pyramidalized along the ν_4 oop bending mode in the excited state, and the potential energy profile along the improper dihedral that breaks the planarity is characterized by a double-well. This scenario prevents the adoption of the harmonic approximation, at least to describe the PES along such degree of freedom. Trofimov *et al.*[58] attributed this feature to a coupling with an O $1s \rightarrow 3s$ state and computed the spectra with a LVC vibronic model. In this case, the effect of vibronic coupling is larger than for C $1s \rightarrow \pi^*$, but it remains moderate as it mainly introduces a broadening, smearing out the main vibronic bands. The parameterization of a full dimensionality LVC model on the grounds of EOM-CCSD data might be done generalizing the strategies we adopted in combination with TD-DFT for some nucleobases,[62–64] but this lies beyond the scope of this work. Therefore, we adopted a simpler strategy and removed the mode ν_4 from our harmonic calculations. This implicitly constrains the equilibrium geometry to C_{2v} symmetry, analogously to what is obtained with a LVC model neglecting the interstate coupling. Our computed spectrum is shown in Fig. 11(a), where vibronic progressions along modes ν_2 and ν_3, both connected to the CO stretching, are clearly visible. Figure 11(b) compares our computed spectrum (red shifted by 1.72 eV) to the experimental one taken from the work of Trofimov *et al.*, showing that our predicted shape is remarkably narrower than the experimental one (and the LVC one computed in the work of Trofimov *et al.*). Similar to what happens for the C $1s \rightarrow \pi^*$ spectrum, the main difference of our

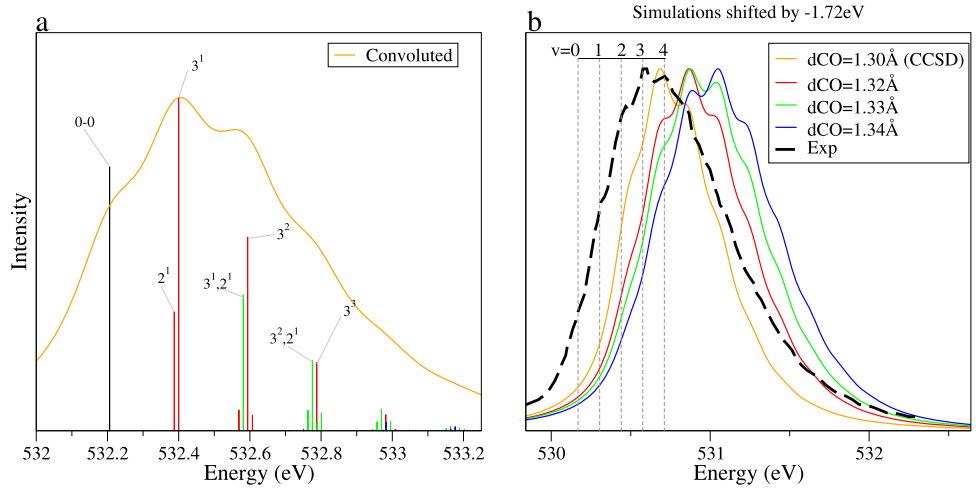


FIG. 11. Formaldehyde. Panel (a): Simulation of the O 1s → π^* band at 0 K, on the grounds of EOM-CCSD PES, adopting a Lorentzian broadening of HWHM = 0.13 eV. The spectra are computed at the C_{2v} stationary point in the excited state, removing the out-of-plane bending which presents a double-well type anharmonic profile in the excited state. Vertical lines show the vibronic transitions, similar to what presented in Fig. 10. Modes 2 and 3 correspond to a combination of C=O stretching and H—C—H bending. Panel (b): Comparison with the experiment. Different computed spectra have been obtained by artificially increasing the CO bond length in the excited state up to 1.34 Å. Computed spectra have been all red-shifted by 1.72 eV. This value was chosen to better superimpose the spectrum directly obtained by CCSD data (i.e., with dCO = 1.30 Å) to the experiment. Vertical dashed lines show the main progression as assigned from the experiment.[58]

calculations with respect to those of Trofimov *et al.* is connected to the underestimation of the elongation of the C=O bond-length—that is, 1.30 Å in our calculations vs ~1.34 Å in their LVC model. This is proven in Fig. 11(b) recomputing our spectrum by artificially increasing the CO bond length of the excited state up to ~1.34 Å and showing that the spectral shape obtained in this case is in fair agreement with the experiment.

The LVC assumption of equal frequencies in the ground and excited states makes the bands much more spaced and therefore resolved, giving overall a not satisfactory agreement with the experiment of Trofimov *et al.* In Ref. 58, Trofimov *et al.* phenomenologically corrected for this error with a modified LVC model, where excited state frequencies were scaled. For example, while the CO stretch frequency is 1751 cm^{-1} in the ground state, it matches the experimental value, 1079 cm^{-1}, in the O 1*s* → π^* state. In principle, our AH model accounts for frequency changes and, in fact, we predict a remarkable decrease in frequency of 260 and 85 cm^{-1} for the two modes ν_2 and ν_3 with contributions from the CO stretch. This goes in the direction of the experimental observation, but the frequency change is underestimated. As a consequence, the spacing of our computed vibronic bands is still too large with respect to the experiment [check the vertical lines plotted in Fig. 11(b)]. This finding is in line with the underestimation of the elongation of the CO bond length. In fact, it is reasonable that the longer the bond becomes the lower its vibrational frequency is. It is also possible that remarkable anharmonic effects play a role in the excited state.

Going back to the broadening induced by progressions along mode ν_4, neglected in our calculations, in Fig. 2 of the supplementary material, we also report a one-dimensional scan of the O 1*s* → π^* along the ν_4 mode in internal coordinates which shows that, at the adopted level of theory, the C_{2v} barrier is quite low (0.3–0.5 cm^{-1}), not supporting any bound state. Moreover, the data in Table XI (supplementary material) indicate that, apart from the value of the oop angle, all other geometrical parameters are practically identical in the C_{2v} transition-state and in the true C_s minimum. In this situation, the energy profile might be approximately substituted by a harmonic curve. We therefore pragmatically checked the possible effect of a frequency change on the ν_4 mode by replacing the double-well profile with a harmonic one in the vibronic calculations, attributing to ν_4 different harmonic frequencies, from ~100 to ~2000 cm^{-1} (see Fig. 3 of the supplementary material). The predicted spectrum changes only slightly, confirming that for this transition, a full LVC treatment including interstate couplings is necessary to capture the effect of ν_4 on the spectrum.

III. CONCLUSIONS

The capability of the core-valence-separated CC methods eCVS-CCS, eCVS-CC2, and eCVS-CCSD to yield reliable core spectra has been tested at the carbon, oxygen, and fluorine K-edges of formaldehyde, formic acid, formyl fluoride, and carbonyl fluoride, with some emphasis on the first 1s → π^* transition as well as the region of Rydberg transitions (in the case of formaldehyde). Different basis sets of double, triple, and quadruple-ζ quality have been tested with augmenting diffuse functions, core polarization functions, and Rydberg functions. While eCVS-CCS, not surprisingly, always fails in describing the core excitations at all K-edges, the eCVS-CC2 approach yields XAS spectra quite similar to eCVS-CCSD at the carbon K-edge, apart from a different systematic shift compared to the experiment. At the oxygen K-edge, the eCVS-CC2 spectra are typically "compressed" compared to the eCVS-CCSD ones, i.e., the higher energy excitations fall too close to the dominant 1s → π^* band. At the fluorine K-edge, eCVS-CC2 and eCVS-CCSD spectra are very different, and using eCVS-CC2 is not advisable.

The eCVS-CCSD scheme is found to yield spectra in satisfactory agreement with the experiment at all edges, once a systematic realignment has been applied and an appropriate basis set is chosen. The eCVS-CCSD IEs are also found in satisfactory agreement with the experiment, though at instances slightly misaligned with respect to the systematic shift of the absorption spectrum, yielding term values which are usually slightly overestimated. The eCVS-CC2 IEs are always too low, yielding too low term values.

We also computed the C $1s \rightarrow \pi^*$ and O $1s \rightarrow \pi^*$ vibrationally resolved spectra of formaldehyde, with the AH harmonic model, on the grounds of EOM-CCSD quadratic expansions of the ground and excited states. The overall agreement with the experiment is reasonable and qualitatively reproduces the observed isotopic effects. Accounting for the differences of the energy Hessians of the initial and final states of the transitions introduces frequency changes and Duschinsky mixings that remarkably improve the spacings of the main vibronic bands, although for the O $1s \rightarrow 3s$ state they are still too large probably due to important anharmonic effects. Although we neglect interstate electronic couplings, which according to Trofimov *et al.*[58,60,61] have some moderate effects on the spectral shapes, most of the inaccuracies of our computed spectra can be traced back to the underestimation of the elongation of the CO bond in the excited states. This indicates that for accurate vibronic line shapes of these core-excitations inclusion of the effect of triple excitations scheme may be necessary.

SUPPLEMENTARY MATERIAL

See the supplementary material for the raw spectral data, the geometrical parameters of the ground and core excited states used in the vibronic analysis of formaldehyde, and the comparison between the AH and AS simulation for the C band.

ACKNOWLEDGMENTS

R.F. and S.C. acknowledge support from DTU Chemistry (start-up post-doc grant) and from the Independent Research Fund Denmark, DFF-Forskningsprojekt2, Grant No. 7014-00258B. This work was financially supported by an ERC advanced Grant XRAY-onACTIVE, No. 340279. L.C. acknowledges the start-up fund from the Johns Hopkins University.

REFERENCES

[1] M. Roemelt and F. Neese, J. Phys. Chem. A **117**, 3069 (2013).
[2] J. Sheehy, T. Gil, C. Winstead, R. Farren, and P. Langhoff, J. Chem. Phys. **91**, 1796 (1989).
[3] N. A. Besley, A. T. Gilbert, and P. M. Gill, J. Chem. Phys. **130**, 124308 (2009).
[4] H. Ågren and H. J. A. Jensen, Chem. Phys. **172**, 45 (1993).
[5] N. A. Besley and F. A. Asmuruf, Phys. Chem. Chem. Phys. **12**, 12024 (2010).
[6] S. Coriani, O. Christiansen, T. Fransson, and P. Norman, Phys. Rev. A **85**, 022507 (2012).
[7] S. Coriani, T. Fransson, O. Christiansen, and P. Norman, J. Chem. Theory Comput. **8**, 1616 (2012).
[8] U. Ekström, P. Norman, V. Carravetta, and H. Ågren, Phys. Rev. Lett. **97**, 143001 (2006).
[9] W. D. Derricotte and F. A. Evangelista, Phys. Chem. Chem. Phys. **17**, 14360 (2015).
[10] P. Verma, W. D. Derricotte, and F. A. Evangelista, J. Chem. Theory Comput. **12**, 144 (2016).
[11] P. Norman and A. Dreuw, Chem. Rev. **118**, 7208 (2018).
[12] M. L. Vidal, X. Feng, E. Epifanovsky, A. I. Krylov, and S. Coriani, J. Chem. Theory Comput. **15**, 3117 (2019).
[13] R. J. Bartlett and M. Musial, Rev. Mod. Phys. **79**, 291 (2007).
[14] T. Helgaker, S. Coriani, P. Jørgensen, K. Kristensen, J. Olsen, and K. Ruud, Chem. Rev. **112**, 543 (2012).
[15] O. Christiansen, P. Jørgensen, and C. Hättig, Int. J. Quantum Chem. **68**, 1 (1998).
[16] O. Christiansen, Theor. Chem. Acc. **116**, 106 (2006).
[17] S. Coriani and H. Koch, J. Chem. Phys. **143**, 181103 (2015).
[18] G. Remmers, M. Domke, A. Puschmann, T. Mandel, C. Xue, G. Kaindl, E. Hudson, and D. Shirley, Phys. Rev. A **46**, 3935 (1992).
[19] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic Structure Theory* (Wiley, 2004).
[20] H. Koch and P. Jørgensen, J. Chem. Phys. **93**, 3333 (1990).
[21] E. R. Davidson, J. Comput. Phys. **17**, 87 (1975).
[22] O. Christiansen, A. Halkier, H. Koch, P. Jørgensen, and T. Helgaker, J. Chem. Phys. **108**, 2801 (1998).
[23] L. S. Cederbaum, W. Domcke, and J. Schirmer, Phys. Rev. A: At., Mol., Opt. Phys. **22**, 206 (1980).
[24] F. Santoro, R. Improta, A. Lami, J. Bloino, and V. Barone, J. Chem. Phys. **126**, 084509 (2007).
[25] F. J. A. Ferrer, J. Cerezo, J. Soto, R. Improta, and F. Santoro, Comput. Theor. Chem. **1040**, 328 (2014).
[26] F. Pawłowski, P. Jørgensen, J. Olsen, F. Hegelund, T. Helgaker, J. Gauss, K. L. Bak, and J. F. Stanton, J. Chem. Phys. **116**, 6482 (2002).
[27] M. Nakata, K. Kohata, T. Fukuyama, K. Kuchitsu, and C. Wilkins, J. Mol. Struct. **68**, 271 (1980).
[28] G. H. Kwei and R. Curl, Jr., J. Chem. Phys. **32**, 1592 (1960).
[29] R. F. Miller and R. Curl, Jr., J. Chem. Phys. **34**, 1847 (1961).
[30] K. Aidas, C. Angeli, K. L. Bak, V. Bakken, R. Bast, L. Boman, O. Christiansen, R. Cimiraglia, S. Coriani, P. Dahle, E. K. Dalskov, U. Ekström, T. Enevoldsen, J. J. Eriksen, P. Ettenhuber, B. Fernández, L. Ferrighi, H. Fliegl, L. Frediani, K. Hald, A. Halkier, C. Hättig, H. Heiberg, T. Helgaker, A. C. Hennum, H. Hettema, E. Hjertenæs, S. Høst, I.-M. Høyvik, M. F. Iozzi, B. Jansik, H. J. A. Jensen, D. Jonsson, P. Jørgensen, J. Kauczor, S. Kirpekar, T. Kjærgaard, W. Klopper, S. Knecht, R. Kobayashi, H. Koch, J. Kongsted, A. Krapp, K. Kristensen, A. Ligabue, O. B. Lutnæs, J. I. Melo, K. V. Mikkelsen, R. H. Myhre, C. Neiss, C. B. Nielsen, P. Norman, J. Olsen, J. M. H. Olsen, A. Osted, M. J. Packer, F. Pawlowski, T. B. Pedersen, P. F. Provasi, S. Reine, Z. Rinkevicius, T. A. Ruden, K. Ruud, V. V. Rybkin, P. Salek, C. C. M. Samson, A. S. de Merás, T. Saue, S. P. A. Sauer, B. Schimmelpfennig, K. Sneskov, A. H. Steindal, K. O. Sylvester-Hvid, P. R. Taylor, A. M. Teale, E. I. Tellgren, D. P. Tew, A. J. Thorvaldsen, L. Thøgersen, O. Vahtras, M. A. Watson, D. J. D. Wilson, M. Ziolkowski, and H. Ågren, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **4**, 269 (2014).
[31] T. H. Dunning and P. J. Hay, *Methods of Electronic Structure Theory* (Springer, 1977), pp. 1–27.
[32] K. Kaufmann, W. Baumeister, and M. Jungen, J. Phys. B: At., Mol. Opt. Phys. **22**, 2223 (1989).
[33] J. F. Stanton, J. Gauss, L. Cheng, M. E. Harding, D. A. Matthews, and P. G. Szalay, "CFOUR, coupled-cluster techniques for computational chemistry, a quantum-chemical program package," With contributions from A. A. Auer, R. J. Bartlett, U. Benedikt, C. Berger, D. E. Bernholdt, Y. J. Bomble, O. Christiansen, F. Engel, R. Faber, M. Heckert, O. Heun, M. Hilgenberg, C. Huber, T.-C. Jagau, D. Jonsson, J. Jusélius, T. Kirsch, K. Klein, W. J. Lauderdale, F. Lipparini, T. Metzroth, L. A. Mück, D. P. O'Neill, D. R. Price, E. Prochnow, C. Puzzarini, K. Ruud, F. Schiffmann, W. Schwalbach, C. Simmons, S. Stopkowicz, A. Tajti, J. Vázquez, F. Wang, J. D. Watts and the integral packages MOLECULE (J. Almlöf and P. R. Taylor), PROPS (P. R. Taylor), ABACUS (T. Helgaker, H. J. Aa. Jensen, P. Jørgensen, and J. Olsen), and ECP routines by A. V. Mitin and C. van Wüllen. For the current version, see http://www.cfour.de.
[34] J. Gauss, J. F. Stanton, and R. J. Bartlett, J. Chem. Phys. **95**, 2623 (1991).
[35] H. Koch, H. J. A. Jensen, P. Jørgensen, T. Helgaker, G. E. Scuseria, and H. F. Schaefer, J. Chem. Phys. **92**, 4924 (1990).
[36] J. Gauss and J. F. Stanton, Chem. Phys. Lett. **276**, 70 (1997).

[37]J. F. Stanton and J. Gauss, Int. Rev. Phys. Chem. **19**, 61 (2000).
[38]J. F. Stanton, J. Chem. Phys. **99**, 8840 (1993).
[39]J. F. Stanton and J. Gauss, Theor. Chim. Acta **91**, 267 (1995).
[40]F. J. A. Ferrer and F. Santoro, Phys. Chem. Chem. Phys. **14**, 13549 (2012).
[41]F. Santoro and J. Cerezo, "FCclasses v. 3.0," Available upon request.
[42]T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989).
[43]N. Sylvetsky, M. K. Kesharwani, and J. M. Martin, J. Chem. Phys. **147**, 134106 (2017).
[44]S. Machado, G. Camiletti, A. C. Neto, F. Jorge, and R. S. Jorge, Mol. Phys. **107**, 1713 (2009).
[45]M. Douglas and N. M. Kroll, Ann. Phys. **82**, 89 (1974).
[46]B. A. Hess, Phys. Rev. A **33**, 3742 (1986).
[47]K. G. Dyall, J. Chem. Phys. **106**, 9618 (1997).
[48]W. Liu and D. Peng, J. Chem. Phys. **131**, 031104 (2009).
[49]M. Tashiro, M. Ehara, H. Fukuzawa, K. Ueda, C. Buth, N. V. Kryzhevoi, and L. S. Cederbaum, J. Chem. Phys. **132**, 184302 (2010); e-print arXiv:1004.3092.
[50]N. A. Besley, M. J. G. Peach, and D. J. Tozer, Phys. Chem. Chem. Phys. **11**, 10350 (2009).
[51]R. H. Myhre, T. J. A. Wolf, L. Cheng, S. Nandi, S. Coriani, M. Gühr, and H. Koch, J. Chem. Phys. **148**, 064106 (2018).
[52]K. J. Oosterbaan, A. F. White, and M. Head-Gordon, J. Chem. Phys. **149**, 044116 (2018).
[53]K. J. Oosterbaan, A. F. White, and M. Head-Gordon, J. Chem. Phys. **149**, 139901 (2018).
[54]O. Christiansen, H. Koch, and P. Jørgensen, Chem. Phys. Lett. **243**, 409 (1995).
[55]H. Koch, O. Christiansen, P. Jørgensen, and J. Olsen, Chem. Phys. Lett. **244**, 75 (1995).
[56]K. C. Prince, R. Richter, M. de Simone, M. Alagia, and M. Coreno, J. Phys. Chem. A **107**, 1955 (2003).
[57]M. Robin, I. Ishii, R. McLaren, and A. Hitchcock, J. Electron Spectrosc. Relat. Phenom. **47**, 53 (1988).
[58]A. B. Trofimov, E. V. Gromov, H. Köppel, J. Schirmer, K. C. Prince, R. Richter, M. D. Simone, and M. Coreno, J. Phys. B: At., Mol. Opt. Phys. **36**, 3805 (2003).
[59]I. Ishii and A. Hitchcock, J. Chem. Phys. **87**, 830 (1987).
[60]A. B. Trofimov, T. E. Moskovskaya, E. V. Gromov, H. Köppel, and J. Schirmer, Phys. Rev. A **64**, 022504 (2001).
[61]A. B. Trofimov, E. V. Gromov, T. E. Moskovskaya, and J. Schirmer, J. Chem. Phys. **113**, 6716 (2000).
[62]D. Picconi, A. Lami, and F. Santoro, J. Chem. Phys. **136**, 244104 (2012).
[63]J. Cerezo, Y. Liu, N. Lin, X. Zhao, R. Improta, and F. Santoro, J. Chem. Theory Comput. **14**, 820 (2018).
[64]Y. Liu, L. Martinez-Fernandez, J. Cerezo, G. Prampolini, R. Improta, and F. Santoro, Chem. Phys. **515**, 452 (2018), special issue on "Ultrafast photoinduced processes in polyatomic molecules: Electronic structure, dynamics and spectroscopy," dedicated to Wolfgang Domcke on the occasion of his 70th birthday.