Inner-shell photoabsorption and photoionisation cross-sections of valence excited states from asymmetric-Lanczos equation-of-motion coupled cluster singles and doubles theory

Moitra, Torsha; Coriani, Sonia; Cabral Tenorio, Bruno Nunes

Published in:
Molecular Physics

Link to article, DOI:
10.1080/00268976.2021.1980235

Publication date:
2021

Document Version
Peer reviewed version

Citation (APA):
Moitra, T., Coriani, S., & Cabral Tenorio, B. N. (2021). Inner-shell photoabsorption and photoionisation cross-sections of valence excited states from asymmetric-Lanczos equation-of-motion coupled cluster singles and doubles theory. *Molecular Physics*, *119*(21-22), Article e1980235. https://doi.org/10.1080/00268976.2021.1980235

Inner-shell photoabsorption and photoionization cross-sections of valence excited states from asymmetric-Lanczos equation-of-motion coupled cluster singles and doubles theory

Torsha Moitra[a], Sonia Coriani[a] and Bruno Nunes Cabral Tenorio[a]

[a]DTU Chemistry–Department of Chemistry, Technical University of Denmark, Kemitorvet Bldg 207, 2800 Kongens Lyngby, Denmark

ARTICLE HISTORY
Compiled September 8, 2021

ABSTRACT
The asymmetric-Lanczos based equation-of-motion coupled cluster formalism to compute photoabsorption and photoionization cross-sections of valence excited states [Tenorio *et al.*, J. Chem. Phys., 151, 184106 (2019)] has been adapted to enable the calculation of photoabsorption and photoionization spectral signatures from inner-shell electrons of such states. Since excited state properties depend on both the electronic character of the molecular wavefunction and on the nuclear dynamics, we computed the photoionization spectra using both ground and excited state optimized geometries. The total cross-section profiles were generated for the first two electronically excited states of water, ammonia, ethylene, and uracil by two different methodologies: an analytic continuation procedure based on the Padé approximants and by the Stieltjes imaging procedure. A comparison with literature results, whenever available, is presented. Remarkable differences were observed between the results of the core-ionization cross sections of the valence excited states yielded by the two quadrature approaches, at variance from previous studies on the valence photoionization cross sections of ground states and of valence excited states. Their origin remains unclear.

KEYWORDS
X-ray spectroscopy, Photoionization, cross-sections, Photo-absorption

1. Introduction

The last decade of advances in synchrotron radiation sources, free-electron lasers and tabletop light sources in the soft X-ray spectral region based on high-harmonic generation have enabled a deeper understanding of dynamical processes induced by the interaction of ultrashort ultra-violet (UV) or X-ray pulses with molecules and solids.

Time-resolved X-ray absorption (TR-XA) spectroscopy is currently a blooming research field with application ranging from isolated molecules to noncrystalline biological systems and solids [1–10]. A large number of theoretical methods have been proposed over the last five years to assist in the interpretation of TR-XAS experiments [4, 11–14]. In the spirit of the work presented by Neville *et al.* [15] for the Algebraic Diagrammatic Construction approach, we here extend our Lanczos-based (equation-of-motion) coupled cluster methodologies [16, 17] to the determination of inner-shell

---

CONTACT Sonia Coriani: soco@kemi.dtu.dk, Bruno N.C. Tenorio: brncat@dtu.dk

(K-edge) photoabsorption and photoionization cross-sections of transient states, that are of interest for the simulation of time-resolved X-ray spectroscopy.

The core excitation cross-sections from ground and valence excited state can be aptly partitioned into two regions with the ionization energy separating the two. The behaviour of the pre- and post-ionization regions are quite different with excitations to bound and continuum final states, respectively. Conventional quantum chemical methods rely on quadratically integrable (L^2) basis sets which describe the discrete region accurately but fail for the continuum part. Extending the use of the L^2 basis set beyond the ionization threshold requires specialized quadrature algorithms.

The Stieltjes Imaging technique and the analytic continuation approach based on Padé approximants are two such methods, used since the 1970s to obtain total photoionization cross-sections [18–22]. They are convenient in that they allow to treat the discrete and continuum part of the spectra on an equal footing, even though their application is not entirely black-box. Moreover, only total photoionization cross-sections can be extracted, due to the lack of proper asymptotic boundary conditions, which hampers the determination of individual channel cross-sections and asymmetry parameters. For the latter, approaches that explicitly solve the Schrödinger equation for the continuum are required, like for instance the multi-centric B-spline (TD-)DFT methods [23–25].

Albeit these limitations, total photoionization cross-sections have been extensively studied in the past two decades using both quadrature methodologies. Modern applications are in particular based on (unconverged) effective ground state and transient excited state pseudo-spectra generated by means of Lanczos eigensolvers [15–17, 26–37]. An attractive characteristic of the Lanczos algorithms is that they yield approximations to the full eigenspectrum, which moreover converges from the extreme ends.

As anticipated, we herein report on the development and application of a core-valence-separated (CVS) asymmetric-Lanczos scheme within the equation-of-motion coupled cluster singles and doubles (EOM-CCSD) method together with the Stieltjes imaging procedure and the analytic continuation procedure based on the Padé approximants to study the K-edge photoionization curves of the first two valence excited states in the molecules water, ammonia, ethylene, and uracil. The cross-sections obtained with both quadrature procedures are critically compared among each other and with other available computed results [15].

2. Theory

2.1. Lanczos CVS-EOM-CCSD: core-excited transition strengths from valence-excited states

Atomic units will be used throughout, unless otherwise specified. Within EOM-CC [38] theory as well as CC response theory [39, 40], the oscillator strength (f_{ij}) between an initial excited state $|i\rangle$ and a final state $|j\rangle$ can be written as

$$f_{ij} = \frac{2}{3}(\omega_j - \omega_i)(\mathcal{S}_{ij}^{xx} + \mathcal{S}_{ij}^{yy} + \mathcal{S}_{ij}^{zz}), \tag{1}$$

where, e.g., $\mathcal{S}_{ij}^{xx}$ is the transition strength for the x Cartesian component of the electric dipole operator d (and similarly for the other components). We follow the notation of Refs. 41, 42 and 43, and write the EOM-CC transition strength as product of EOM-CC transition moments, $\mathcal{S}^{xx}_{ij} = \mathcal{T}^{x}_{ij}\mathcal{T}^{x}_{ji}$, with

$$\mathcal{T}^{x}_{ij} = (L^i \mathbf{A}^x R^j) - (L^i{\cdot}\xi^x)(\bar{t}{\cdot}R^j) - (L^i{\cdot}R^j)(\bar{t}{\cdot}\xi^x) \quad . \tag{2}$$

In Eq. (2), L^i and R^j are left and right excitation vectors, respectively; $\bar{t}$ indicates the ground state Lagrangian multipliers, and ξ^x is the CC property gradient vector for the electric dipole operator component d_x [39].

According to both the EOM-CC framework [38, 44] and to CC linear response [39, 40], the right and left excitation vectors and their excitation energies (ω_i, ω_j) can be obtained as eigenvectors and eigenvalues of the CC Jacobian matrix $\mathbf{A}$

$$L^j\mathbf{A} = \omega_j L^j; \qquad \mathbf{A}R^j = \omega_j R^j; \tag{3}$$

under the biorthogonality condition $L^j R^i = \delta_{ji}$. The Jacobian matrix $\mathbf{A}$ is defined as

$$A_{\mu\nu} = \langle\mu|\exp(-T)[H,\tau_\nu]\exp(T)|\mathrm{HF}\rangle \equiv \langle\mu|[\overline{H},\tau_\nu]|\mathrm{HF}\rangle \tag{4}$$

where we also introduced the similarity transformed Hamilton operator, $\overline{H} = \exp(-T)H\exp(T)$. The cluster operator T is, as usual, a sum of excitation operators, τ_μ, each weighted by its corresponding amplitude t_μ, $T = \sum_\mu t_\mu\tau_\mu$. The excited Slater determinant $|\mu\rangle$ is generated applying the excitation operator τ_μ to the reference Hartree-Fock state, $|\mu\rangle = \tau_\mu\,|\mathrm{HF}\rangle$ [45].

The EOM-CC property Jacobian matrix $\mathbf{A}^x$ in Eq. (2) can be written as

$$A^x_{\mu\nu} = \langle\mu|\overline{d}_x|\nu\rangle - \delta_{\mu\nu}\langle\mathrm{HF}|\overline{d}_x|\mathrm{HF}\rangle \tag{5}$$

As also pointed out elsewhere [43], $A^x_{\mu\nu}$ is different from the one of CC response theory,

$$A^{x,\mathrm{LR}}_{\mu\nu} = \langle\mu|\exp(-T)[d_x,\tau_\nu]\exp(T)|\mathrm{HF}\rangle \equiv \langle\mu|[\overline{d}_x,\tau_\nu]|\mathrm{HF}\rangle \ , \tag{6}$$

but it can be easily obtained from it with minor modifications [43]

$$\mathbf{A}^x_{\mu\nu} = \mathbf{A}^{x,\mathrm{LR}}_{\mu\nu} + \langle\mu|\,\tau_\nu\overline{d}_x\,|\mathrm{HF}\rangle\,(1-\delta_{\mu\nu}) \tag{7}$$

Eq. (3) can be solved using different types of solvers, including the asymmetric Lanczos algorithm, as extensively explored for instance by Coriani and coworkers [16, 46, 47]. In the CC asymmetric Lanczos algorithm, [46, 47], a truncated tridiagonal representation $\mathbf{T}$ of the Jacobian matrix $\mathbf{A}$ is built and diagonalized, yielding a discretized pseudo-spectrum covering the whole frequency range. The tridiagonal matrix $\mathbf{T}$ is conveniently truncated to a dimension (called *chain length*) $k \ll n$, where n is the full dimension of the excitation space. The non-zero elements of the tridiagonal matrix $\mathbf{T} = \mathbf{P}^T\mathbf{A}\mathbf{Q}$, where $\mathbf{P}^T\mathbf{Q} = \mathbf{1}$, are obtained as

$$\begin{aligned} T_{ll} &\equiv \alpha_l = \mathbf{p}_l^T\mathbf{A}\mathbf{q}_l; \quad T_{l+1,l} \equiv \beta_l = \sqrt{\mathbf{p}_{l+l}^T\mathbf{q}_{l+1}}; \\ T_{l,l+1} &\equiv \gamma_l = \mathrm{sgn}\{\mathbf{p}_{l+1}^T\mathbf{q}_{l+1}\}\beta_l \end{aligned} \tag{8}$$

with

$$\begin{aligned}\mathbf{q}_{l+1} &= \beta_l^{-1}(\mathbf{A}\mathbf{q}_l - \gamma_{l-1}\mathbf{q}_{l-1} - \alpha_l\mathbf{q}_l)\ ; \\ \mathbf{p}_{l+1}^T &= \gamma_l^{-1}(\mathbf{p}_l^T\mathbf{A} - \beta_{l-1}\mathbf{p}_{l-1}^T - \alpha_l\mathbf{p}_l^T)\end{aligned} \tag{9}$$

The vectors $\mathbf{p}_i$ and $\mathbf{q}_i$ are columns of the (rectangular, $n \times k$) matrices $\mathbf{P}$ and $\mathbf{Q}$, respectively. The effective pseudo-spectrum generated diagonalizing $\mathbf{T}$ is known to converge from the extremities towards the exact excitation spectrum [46–48].

We generate the pseudo-spectrum only in the symmetry space of the dipole-allowed final (core-excited) state $|j\rangle$ from the initial valence-excited state $|i\rangle$, see Eq. (2), by using the following normalized vectors as Lanczos seeds

$$\begin{aligned}\mathbf{q}_1 = \mathbf{Q}\mathbf{e}_1 &= {}^ix_q^{-1}\big[\mathbf{A}^x R^i - (\bar{t}\cdot R^i)\xi^x\big] \\ \mathbf{p}_1^T = \mathbf{e}_1^T\mathbf{P}^T &= {}^ix_p^{-1}\big[L^i\mathbf{A}^x - (L^i\cdot\xi^x)\bar{t}\ \big]\end{aligned} \tag{10}$$

The scalars ${}^ix_q$ and ${}^ix_p$ are the normalization factors and $\mathbf{e}_m$ is a vector of length k whose m^{th} element is equal to 1 and all other elements are set to zero. The left superscript i in the normalization factors highlights the implicitly dependence of the seeds on the initial excited state.

The above definition of the Lanczos seeds is based on Eq. (2), where the last term was omitted since it eventually does not contribute to the transition moments due to the biorthogonality condition between excited states. Now, we can straightforwardly calculate the EOM-CCSD transition strengths between the final core excited states (labelled J) and the initial valence excited states (labelled i) via Eq. (2),

$$\begin{aligned}\mathcal{T}^x_{iJ}\cdot\mathcal{T}^x_{Ji} &= ({}^ix_q L^J\mathbf{q}_1)\cdot({}^ix_p\mathbf{p}_1^T R^J) = ({}^ix_q L_J^{(k)}\mathbf{P}^T\mathbf{Q}\mathbf{e}_1)({}^ix_p\mathbf{e}_1^T\mathbf{P}^T\mathbf{Q}R_J^{(k)}) \\ \mathcal{S}^{xx}_{iJ} &= {}^ix_p\,{}^ix_q R_{J,1}^{(k)} L_{J,1}^{(k)}\end{aligned} \tag{11}$$

where we used the back-transformations of the Lanczos eigenvectors $L_J^{(k)}$ and $R_J^{(k)}$ into full space, $L^J = L_J^{(k)}\mathbf{P}^T$ and $R^J = \mathbf{Q}R_J^{(k)}$, and the orthogonality condition $\mathbf{P}^T\mathbf{Q} = \mathbf{1}$. Note once more the use of Capital letters to label the inner-shell states.

The EOM-CCSD valence and core states are obtained using a dual CVS scheme [49, 50]. A "core-only" CVS projector is used to remove the continuum sea of valence ionizations from the core excitation region. This procedure yields a Lanczos pseudo-spectrum starting at the lowest core excitation energy of the chosen edge. A complementary "valence-only" CVS projector is conversely used to ensure the initial valence excited state is orthogonal to the core states.

To obtain final Lanczos states in the core region, the asymmetric Lanczos algorithm is modified to perform the 'core-only' CVS by applying, to both the $\mathbf{p}_i^T$ and $\mathbf{q}_i$ vectors and their linear transformations $\mathbf{p}_i^T\mathbf{A}$ and $\mathbf{A}\mathbf{q}_i$, below generically indicated as vector $\mathbf{b}$, a projector $\mathcal{P}_L^{\div\text{val}}$

$$\begin{cases}\mathcal{P}_L^{\div\text{val}}\ b_l^a &= 0 \qquad\qquad \text{if } l \neq L, \\ \mathcal{P}_L^{\div\text{val}}\ b_{lm}^{ab} &= 0 \qquad \text{if } l \neq L \text{ or } m \neq L\end{cases} \tag{12}$$

at each iteration. Labels a and b indicate here virtual orbitals and L, l, m occupied orbitals. The projector removes all vector elements not referencing to at least one occupied core orbital L or a set of selected core orbitals $\{L\}$, effectively decoupling the core edge from the pure valence excitations. Diagonalization of the resulting CVS-projected $\mathbf{T}$ matrix yields the core excited states as lowest roots [16]. The initial valence excited states are kept orthogonal to the final core states by using the 'valence-only' CVS projector $\mathcal{P}_L^{\div\text{core}}$ during the determination of the valence excited states. The valence-only projector removes the core excitations from the valence excited-state manifold.

Our procedure for calculating the inner-shell pseudo-spectrum from a valence excited state within the EOM-CCSD framework can thus be summarized as follows

(1) We compute the selected initial excitation vectors (R^i and L^i) via a modified Davidson diagonalization of Eq. (3), where a 'valence-only' CVS projector is applied to remove the core excitations from the manifold.
(2) Following that, we generate the Lanczos seeds according to Eq. (10). They are dependent on the initial excited state $|i\rangle$ and ensure that only the space of final core-excited states $|J\rangle$ that are dipole-allowed when accessed from the chosen initial excited state is spanned.
(3) Next, we iteratively build the tridiagonal matrix $\mathbf{T}$, applying at each iteration the core-only CVS projector for the selected K-edge
(4) We then diagonalize $\mathbf{T}$, conveniently truncated to a dimension k, to obtain the final space pseudo-spectrum.
(5) Finally, the oscillator strengths between the initial valence excited state $|i\rangle$ and the final core excited states $|J\rangle$ are computed from Eq. (11).

2.2. Photoionization cross-section from Stieltjes Imaging

Total photoionization cross-sections can be obtained by the Stieltjes Imaging procedure, suggested by Langhoff [18, 19]. The cross-section $\sigma(\omega)$ is the probability for a photon of energy ω to be absorbed by a molecule, and is proportional to the (differential) oscillator strength $f(\omega)$, $\sigma(\omega) \propto f(\omega)$. The oscillator strength $f(\omega)$ is dimensionless.

Below the ionization limit, the oscillator strength corresponds to the usual discrete strengths f_i [51]. In the case of photoionization, the discrete oscillator strengths f_i are replaced by a continuous oscillator strength function $f(\omega)$, which can be calculated, in an analogous way as the discrete ones, from the integral between initial bound state, and a final continuum state $\phi(\omega)$. The latter is a continuum solution of the electronic Schrödinger equation with an energy ω above the ground state [51]. Unlike bound-state wavefunctions, the continuum wavefunctions are not square-integrable, i.e., the integral $\langle\phi(\omega)|\phi(\omega)\rangle$ is infinite.

The *moments* $s(\kappa)$ of the (differential) oscillator strength $f(\omega)$ can be expressed as a sum of discrete (below the ionization limit) and continuum (above the ionization limit) contributions [51],

$$s(\kappa) = \sum_{l=1}^{\text{discr}} \omega_l^\kappa f_l + \int_{\omega_T}^{\infty} \omega^\kappa f(\omega) d\omega \equiv \sum_{l=1}^{\text{discr}} \omega_l^\kappa f_l + S(\kappa) \quad . \tag{13}$$

Similar to the Padé analytic continuation procedure, the Stieltjes Imaging procedure also exploits a discrete electronic pseudo-spectrum obtained from a L^2 basis set calculation. In Stieltjes Imaging, the moments $S(\kappa)$ of the continuum part of the oscillator strength are approximated by a sum of N discrete pseudo-states of the continuum,

$$S(\kappa) = \int_{\omega_T}^{\infty} \omega^{\kappa} f(\omega) d\omega \approx \sum_{i=1}^{N} \bar{\omega}_i^{\kappa} \bar{f}_i \quad . \tag{14}$$

The purpose of the Stieltjes Imaging procedure is to obtain a smoothed representation of the continuum part of the spectrum by computing a set of M "primitive" excitation energies and oscillator strengths that reproduce the sum rules of the oscillator strength. First, one computes a number $(2r)$ of spectral moments $S(-\kappa)$ of the primitive spectrum

$$S(-\kappa) = \sum_{l=1}^{M} f_l \omega_l^{-\kappa}; \qquad \kappa = 0, 1, \cdots, 2r-1 \ . \tag{15}$$

Then, one generates the discretized spectra of order $m = 2, \cdots, r$, called *principal* pseudo-spectra, as generalized (Gaussian) quadrature points (abscissae) $\omega_l^{(m)}$ and weights $f_l^{(m)}$

$$S(-\kappa) = \sum_{l=1}^{m} f_l^{(m)} (\omega_l^{(m)})^{-\kappa}; \qquad \kappa = 0, ..., 2m-1 \tag{16}$$

Finally, one numerically reconstructs the photoionization cross-section spectrum from the Gaussian quadrature points and weights as the *Stieltjes derivatives* of the general discretized principal pseudo-spectra. For a detailed description of the Stieltjes Imaging procedure, we refer the reader to, e.g., Ref. 51.

Based on our previous work [16, 29, 30], we computed the *principal* pseudo-spectra with r ranging from 2 to 20 and only plot the points that are well converged. As in Ref. 17, we used the Lanczos EOM-CCSD transition energies ω_{iJ} and oscillator strengths f_{iJ} (from valence-excited state i to core-excited states J) as primitive set in the Stieltjes Imaging procedure to obtain the cross-section points $\sigma_i(\omega)$ for the initial excited state.

2.3. Photoionization cross-sections from Padé approximants

Following the procedure introduced in Ref. [17], we also compute the inner-shell photoionization cross-section of valence excited state $|i\rangle$, $\sigma_i(\omega)$, from the imaginary part of the averaged dynamic dipole polarizability function of the excited electronic state,

$$\sigma_i(\omega) = \frac{4\pi\omega}{c} \lim_{\eta \to 0} \mathrm{Im}[\overline{\alpha}_i(\omega + i\eta)] \quad . \tag{17}$$

The averaged dynamic dipole polarizability function $\overline{\alpha}_i(z)$ of the excited state $|i\rangle$, is, in turn, approximated by a finite sum

$$\overline{\alpha}_i(z) = \sum_{J=1}^{K} \frac{f_{iJ}}{\omega_{iJ}^2 - z^2} \tag{18}$$

over the set of Lanczos EOM-CCSD transition energies $\{\omega_{iJ}\}$ and oscillator strengths $\{f_{iJ}\}$ between excited states $|i\rangle$ and $|J\rangle$; z is the complex frequency. As before, we use capital letters J and K to indicate core excited states.

Thus, using the expression for the polarizability function in Eq. (18), we calculate $\alpha_i(z)$ at a number of points in the complex plane. These complex points are selected according to the recipe previously described in Ref. 32. The polarizability values at these complex points are subsequently fitted by the Padé approximants providing an analytical representation of $\alpha_i(z)$ in the complex plane. The imaginary part of $\alpha_i(z)$ on the real axis yields the cross-section according to Eq. (17).

3. Computational details

The asymmetric-Lanczos based algorithm for core excitations of valence excited states has been implemented in Dalton [52–54]. Ground-state geometry optimizations were performed at the MP2/cc-pCVTZ level using the CFOUR program [55]. The two lowest-lying excited states' geometries of H_2O and NH_3 were optimized at the EOM-CCSD/aug-cc-pVTZ level using Q-Chem [56]. The excited state geometries of uracil are the EOM-CCSD optimized ones taken of Ref. 11. For ethylene, in addition to the Frank-Condon geometry, we considered a "twisted-pyramidized" (Tw-py) geometric structure, with an 87 degrees dihedral angle HCCH, and one pyramidized C atom. We utilized the values of the geometric parameters from Ref. 57.

Ground and excited state ionization energies were computed at EOM-CCSD/aug-cc-pVTZ level using Q-Chem [56]. We used the CVS scheme applied to the asymmetric Lanczos procedure for the calculation of EOM-CCSD spectral moments as implemented in the Dalton program package [52] to obtain the core-level pseudo-spectra from valence excited state. Chain-length subspaces of $k = 500$ were used. The Dunning basis sets [58–60] were augmented with continuum-like functions, generated according to the prescription of Kaufmann, Baumeister, and Jungen (KBJ) [61], placed at the center of mass. The aug-cc-pVTZ set was used for small molecules (H_2O, NH_3 and C_2H_4) and a combination of aug-cc-pCVDZ (on edge atom) and aug-cc-pVDZ (on other atoms) was used for uracil. The KBJ set contained s, p and d types functions with quantum number ranging from 2 to 8. The pseudo-spectra were then fed into the Stieltjes Imaging procedure and the Padé approximants analytic continuation procedure to generate the continuum part of the cross-section profiles. Our stand-alone Fortran90 codes for Stieltjes Imaging and the analytic continuation procedure were used [16, 29].

4. Results and discussion

4.1. Ionization Energies

The (first) core ionization energy separates the bound and continuum part of the total cross-section profile. For the ground state, ionization energies can be prompty obtained using the EOM-IP-CCSD scheme or, equivalently, the EOM-EE-CCSD scheme restricting the excitation to a "superdiffuse" orbital [50, 62, 63]. On the other hand, it is a bit more problematic to decide where the core ionization boundary for a valence excited state should be and how to accurately compute it. Starting from the electronic configuration of the valence excited state, different final configurations with a hole in

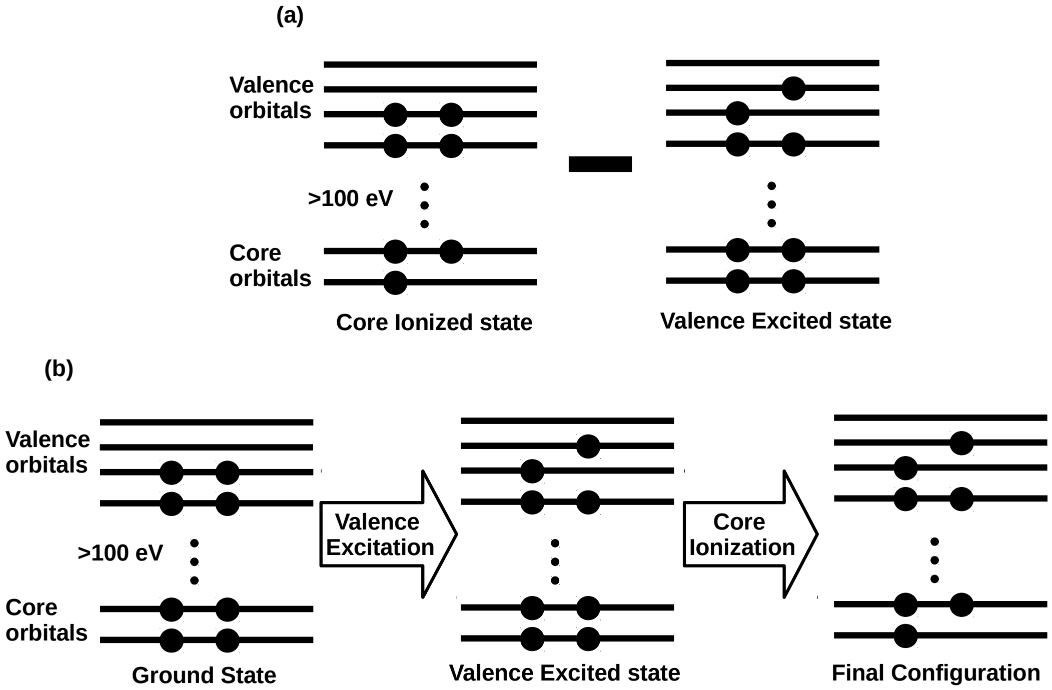

Figure 1.: Calculation of a core ionized state from a valence excited state by: (a) energy difference method (b) using the MOM-CCSD approach to compute the excited state wavefunction, followed by EOM-IP-CCSD core ionization.

the core may be reached. By a shake-down process, one can return to the electronic configuration corresponding to a core ionization from the ground state. The ionization energy of this process is typically obtained by subtracting the valence excitation energy from the ground state core ionization energy [15], as done (to our understanding) in Ref. 15. We will refer to these values as Δ values in the following, e.g. in Table 1. This energy difference method always produces a lower ionization energy for the valence excited states in comparison to the one of the ground state.

Another possible outcome of the core ionization process is a ($2h$-$1p$) final configuration with a hole in the core plus the $1h$-$1p$ valence excitation. One possible way to reach this $2h$-$1p$ state is to employ the Maximum Overlap Method (MOM) together with CCSD [64–66]. Using MOM, one first determines the valence-excited-state self-consistent solutions by pre-defining the set of occupied orbitals that mimic the dominant electronic configuration of the excited state of interest. One then directly computes the energy of the core ionizations from these excited states using EOM-IP. Here we explored two different strategies for the MOM step, namely describing the valence excited state configuration to be of either high spin (triplet), as proposed in Ref. 67, 68, or of low spin (singlet), as done for instance in Ref. 66. The shake-down configurations are also obtained using MOM-CCSD. Either choice has drawbacks. The valence excited state we aimed at probing is a singlet and not a triplet, but using a high-spin reference of similar electronic character ensures that we have a spin-complete wavefunction. The low-spin singlet open shell configuration is prone to variational collapse and also suffers from spin incompleteness. Furthermore, the open-shell configuration leaves us with the choice of ionizing the α or β core electron, which are no longer degenerate (unrestricted case) [67, 68].

Another way to reach the $2h$-$1p$ configuration is as a shake-up of a core ionization of the ground state, and scrutinize therefore the EOM-IP-CCSD results from the ground state to identify this configuration. In either of the last two cases, it is reasonable to expect that core ionization energy for the valence excited state will be somewhat higher than the one for the first core ionization from the ground state.

We note nonetheless that our main focus was not to obtain a very accurate ionization energy, but simply use these estimates as a wall between the excitation and ionization regions of the spectra. We report the ground and excited state ionization energies in Table 1.

Method	Ground state		S_1		S_2		Tw-py
	@S_0	Exp	@S_0	@S_1	@S_0	@S_2	@Tw-py
			$\mathrm{H_2}\mathbf{O}$				
Singlet α	540.66	539.7	548.30	544.14	549.33	544.81	
Triplet α			549.45	545.49	550.51	546.28	
Triplet β			547.80	544.09	548.90	544.69	
Δ			535.05	534.23	531.28	537.21	
			$\mathbf{N}\mathrm{H_3}$				
Singlet α	406.11	405.6	412.72	411.27	413.77	413.44	
Triplet α			413.55	412.96	414.71	414.47	
Triplet β			412.12	411.52	413.36	413.14	
Δ			399.47	396.64	397.91	396.49	
			$\mathbf{C}_2\mathrm{H_4}$				
Singlet α	291.23	290.8	296.32		291.38		295.05
	291.18		296.28		291.36		296.57
Triplet α			296.86		292.30		292.06
			296.80		292.25		292.49
Triplet β			296.18		291.47		291.30
			296.13		291.43		291.69
Δ			283.78		283.17		287.03
			283.73		283.12		288.19
			$\mathbf{C}_4\mathrm{H_4N_2O_2}$				
Singlet α	292.75	291.0	292.48	292.73	292.72	292.91	
	294.84	292.8	293.48	293.87	294.15	295.39	
	296.34	294.4	295.51	295.61	295.23	296.32	
	297.38	295.4	297.04	297.03	297.61	297.47	
Triplet α			292.58	292.78	294.03	292.70	
			293.45	293.85	294.32	293.77	
			295.38	295.52	295.91	295.74	
			297.03	297.02	297.29	297.09	
Triplet β			292.77	293.02	293.47	292.94	
			293.01	293.35	293.72	293.21	
			294.94	295.10	296.00	295.38	
			297.04	297.03	297.31	297.10	
Δ			287.58	289.337	287.18	288.41	
			289.67	291.46	289.27	290.88	
			291.17	292.35	290.77	291.81	
			292.21	293.55	291.81	292.97	
			$\mathrm{C_4H_4}\mathbf{N}_2\mathrm{O_2}$				

Method	Ground state		S_1		S_2		Tw-py
	@S_0	Exp	@S_0	@S_1	@S_0	@S_2	@Tw-py
Singlet α	408.97	406.5	408.57	408.75	408.63	409.14	
	409.43	406.9	408.98	408.96	410.10	409.58	
Triplet α			408.54	408.74	408.67	408.48	
			409.03	408.99	409.55	409.28	
Triplet β			408.48	408.68	408.66	408.40	
			408.89	408.88	409.42	409.16	
Δ			403.80	405.27	403.40	404.63	
			403.40	405.82	403.86	404.94	
			$C_4H_4N_2\mathbf{O}_2$				
Singlet α	541.08	537.6	540.93	540.87	540.32	540.94	
	541.27		544.86	544.88	541.83	545.13	
Triplet α			540.91	540.86	541.37	540.93	
			545.89	546.06	541.50	546.20	
Triplet β			540.90	540.85	541.26	540.90	
			544.84	544.89	541.33	545.12	
Δ			535.91	536.64	535.51	536.38	
			536.10	537.40	535.70	536.81	

Table 1.: Ground and excited-state ionization energies, computed using EOM-CCSD/aug-cc-pVTZ for the smaller molecules and aug-cc-pCVDZ(edge atom)+aug-cc-pVDZ(other atoms) for uracil. The symbols @S_0, @S_1 and @S_2 indicate the geometry at which the results have been obtained: @S_0 is at the Franck Condon point, @S_1 at the first valence excited state geometry and @S_2 is at the second valence excited state geometry. The relevant edge atoms are indicated in bold. All energies are in eV. Experimental results are taken from Refs. 69, 70, 71 and 72.

4.2. Water

The two lowest valence electronic excited states of water correspond to the $^1B_1$ $n3s$ (S_1) and $^1A_2$ $n3p$ (S_2) states, with EOM-CCSD vertical excitation energies obtained at 7.61 eV and 9.38 eV, respectively [17]. A crucial change in geometry is observed upon excitation of water to its low-lying excited states, which is due to the dissociative character of the excited states [74] (see supplemental information, Table S1).

In Figure 2, the ground and excited state pre-edge XA spectra of water are presented. The pre-edge XA of both excited states is dominated by a single intense peak shifted towards the lower energy region with respect to the ground state. The primary peak is characterised as $1s \rightarrow n$(SOMO) excitation in both cases, with the higher energy SOMO being the $3s$ and $3p$ orbital, for S_1 and S_2 states, respectively. This has been schematically presented in Fig. 4. We compute the excited state absorption spectra as the difference in energy between the ground state core excitation and the valence excitation. In the spectra, the characteristic $1s \rightarrow n$(SOMO) transitions of the valence excited states basically overlap energetically, despite the large difference in valence excitation energies, since the two final core excited states of the ground state differ by almost the same amounts, so the differences cancel out. We also observe that the XAS of these excited states are not substantially sensitive to the changes in geometry. Note that, as discussed elsewhere [12, 66], spectral features due to excitations of double or higher character are probably missing in the EOM-CCSD spectra of the

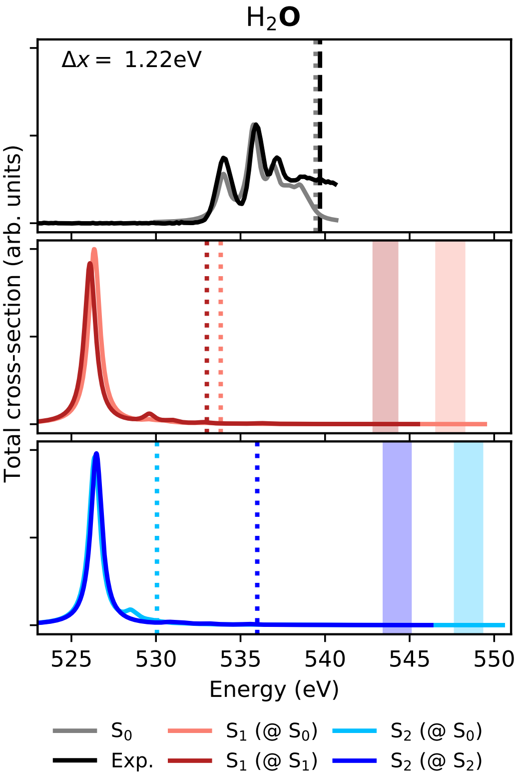

Figure 2.: H_2O: Ground and excited state pre-edge XA spectra obtained with chain length $k = 500$ at EOM-CCSD/aug-cc-pVTZ+KBJ$(7s7p7d)_{n=2-8}$ level, computed at Franck-Condon geometry (label @S_0) and at the geometry of the excited states (labels @S_1 and @S_2). The sticks spectra are convoluted with a Lorentzian broadening of HWHM $= 0.4$ eV. With reference to the values tabulated in Table 1, the dashed vertical lines are the Δ ionization energies obtained as difference between the GS ionization energy and energy of the valence excited state. The shaded vertical strips refer to the ionization energies generated using MOM approach. The experimental results are from Ref. 73. We will follow the same color and style convention throughout, unless otherwise mentioned.

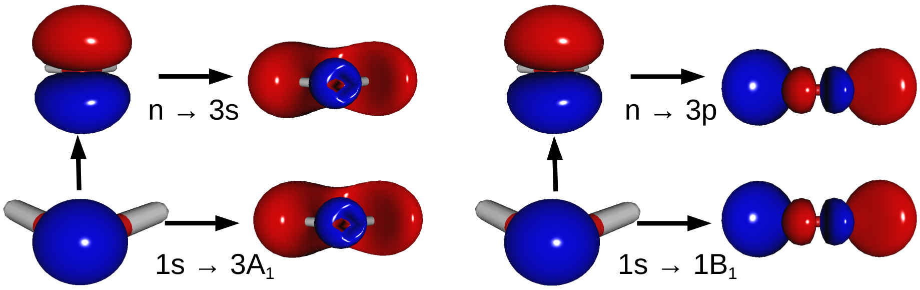

Figure 3.: H_2O: Natural transition orbitals corresponding to the two valence excited states (first row), and to the first two core excitations of the ground state (second row). The 1s → SOMO transition is (1s → n) for both excited states.

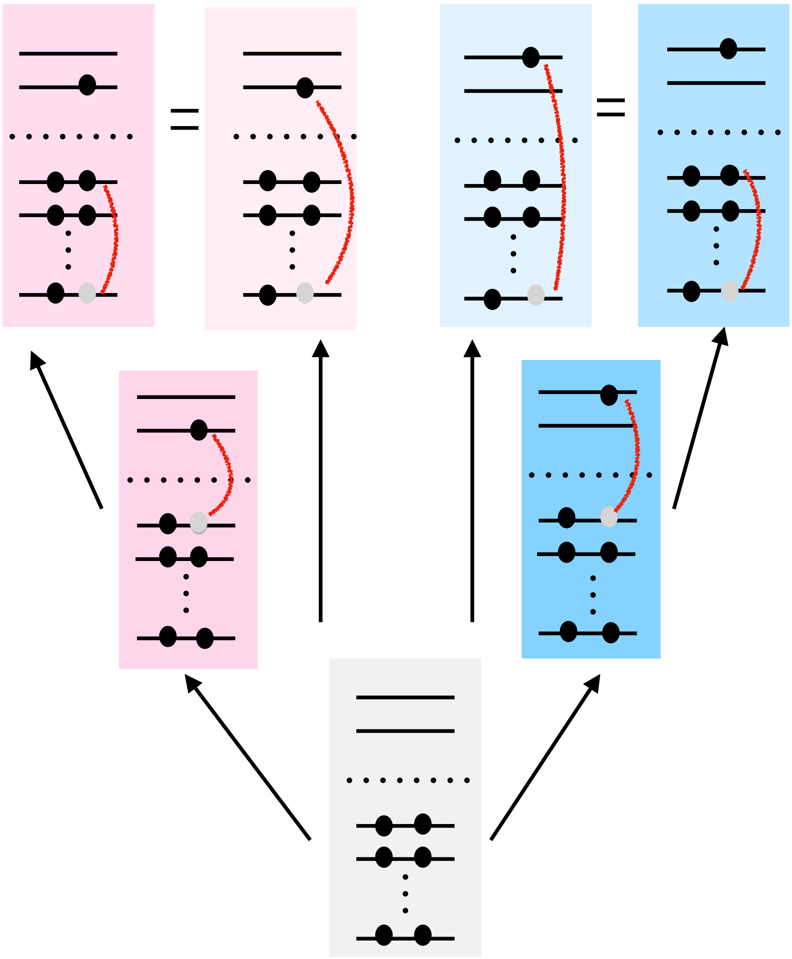

Figure 4.: H_2O. Schematic of the valence and core excitations behind the main features in the XAS spectra (bound part) of the ground and excited states.

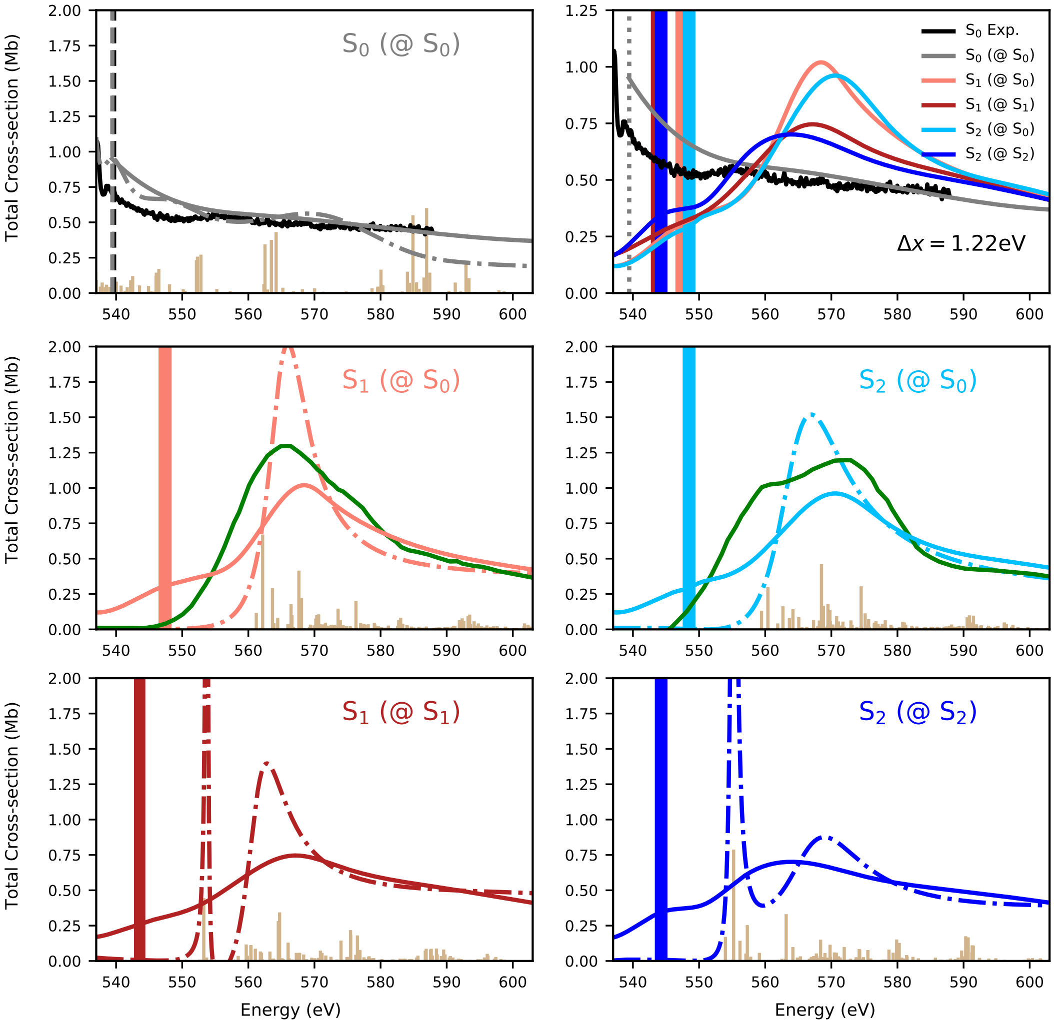

Figure 5.: H_2O: Ground and excited state oxygen K-edge total cross-sections, obtained at the EOM-CCSD/aug-cc-pVTZ+KBJ$(7s7p7d)_{n=2-8}$ level with chain length $k = 500$. The electronic states are color coded: the solid line represents cross-sections obtained using Stieltjes Imaging, and the dash-dot line those from the Padé approximant method. The vertical shaded bars indicate the range of ionization energies as tabulated in Table 1. The computed spectra are shifted by $\Delta x = 1.22$ eV to match the position of the first excitation peak of the ground state (see Fig. 2). Green solid lines are the CVS-ADC(2)-x results from Ref. 15. A composite ground and excited state cross-section obtained from Stieltjes Imaging is plotted in the top right panel. Experimental results are taken from Ref. 73. We will follow the same color and style convention throughout, unless otherwise mentioned.

valence excited states.

The calculated X-ray photoionization cross-sections for the ${}^1B_1/n3s$ (S_1) and ${}^1A_2/n3p$ (S_2) states along with the one for the ground state are shown in Figure 5. The details of the ground state photoionization cross-section calculations are described in our previous work [16]. Similar to what observed by Neville *et al.* [15], both excited-state photoionization cross-sections show a relatively sharp peak centered at around

570 eV, in contrast to the flatter and less intense continuum part of the ground state XA cross-section. It is also evident that the Padé approach gives much sharper and intense peaks than the Stieltjes Imaging procedure. In particular, the cross-sections generated using the Padé approach computed at the excited state optimized geometries show an intense sharp peak at around 555 eV, in excess of the broader peak at around 570 eV observed at the Franck-Condon geometry. This additional feature at the excited state geometries in the Padè case clearly originates from the cluster of intense sticks obtained inbetween 550 eV and 560 eV, which are smoothed away by Stieltjes Imaging. This indicates, in our opinion, a numerical instability in the Padè approximant procedure. A common observation from both Padé and Stieltjes Imaging methods is that the intensity of the peak of the continuum decreases from the Franck-Condon geometry to the optimized excited state geometry. Comparing our results to the CVS-ADC(2)-x ones from Ref. 15 at the Franck-Condon geometry, we note overall qualitative agreement, with the CVS-ADC(2)-x bands being broader and redshifted with respect to ours Stieltjes for S_1 and blueshifted for S_2.

The upper right panel of Figure 5 shows a comparison of the ground state and two excited states' cross-sections from Stieltjes Imaging. As seen, the two bands overlap, with the one of the S_1 state being more prominent and their maxima being roughly 5-7 eV apart. It remains unclear whether the two excited states could be discriminated experimentally from the measurement of the total photoionization cross-sections.

4.3. **Ammonia**

The calculated XAS pre- and post-edge cross-sections for the ground and the ${}^1A_2''/n3s$ and ${}^1E''/n3p_e$ states of NH_3 are shown in Figure 6 and 7, respectively. The (frozen-core) EOM-CCSD energies for these two excited states at the Franck-Condom geometry are 6.65 eV and 8.20 eV. Notice that the two excited states have planar structure at their respective minima [75], with a N-H bond distance of 1.044 Å and 1.016 Å for S_1 and S_2, respectively. The photoabsorption cross-section of the ground state presents multiple peaks below the IE, which have been characterized in previous work, see e.g. Refs. 15, 46, 76. In contrast, XA spectra of the two excited states are characterized by one dominant single peak at about the same excitation energy (see Fig. 6). Similar to water, these peaks predominantly correspond to an excitation from the $1s_\mathrm{N}$ orbital to the singly occupied n orbital (the hole orbital of the valence excited state). The difference in energy between the two final core excited states is similar to the difference in energy between the two valence states, so that the characteristic $1s \rightarrow$ n transitions are practically overlapping, as it happened in water. The pre-edge XAS from the S_2 state presents a shoulder to the main peak at about 1.5 eV higher energy.

Above the ionization threshold of the ground state, the photoionization cross-section is seen to gradually decrease with basically no structure. We refer to Ref. 16 for additional details of the calculation of the ground state XA photoionization cross-section. Conversely, the continuum region of the valence excited states shows broad peaks when using Stieltjes Imaging and (unphysical) sharp features using the Padé approximant method. As reported for water, the cross-sections of ammonia computed using Padé approximants are higher and narrower than those obtained using Stieltjes Imaging. Also, the Stieltjes cross-sections computed with CVS-ADC(2)-x and at the Franck-Condon geometry [15] show a peak in the continuum which is slightly red-shifted with respect to the CVS-EOM-CCSD peak.

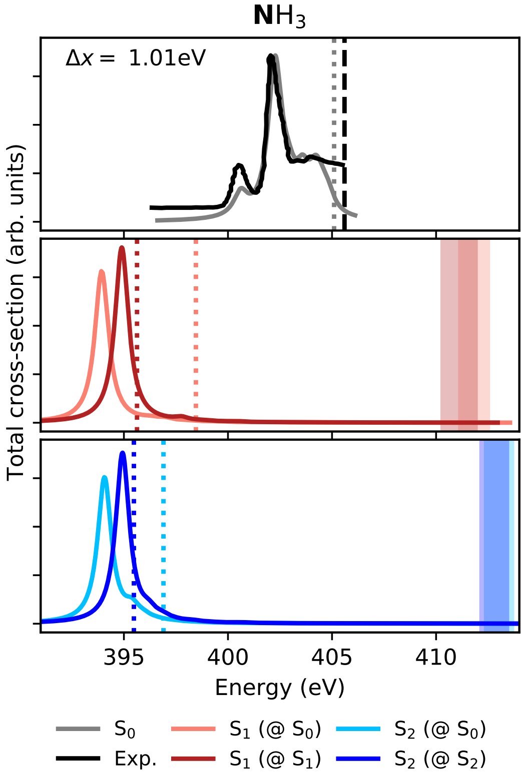

Figure 6.: NH$_3$. Ground and excited state pre-edge XA spectra obtained with chain length $k = 500$ at EOM-CCSD/aug-cc-pVTZ+KBJ$(7s7p7d)_{n=2-8}$ level, computed at Franck-Condon geometry (label @S$_0$) and at the geometry of the excited states (labels @S$_1$ and @S$_2$). The sticks spectra are convoluted with a Lorentzian broadening of HWHM = 0.4 eV. The computed spectra are shifted by $\Delta x = 1.01$ eV to match the position of the first ground-state excitation peak. For more details on color code and styles used here, see caption of Fig. 2. Experimental result taken from Ref. 73.

4.4. ***Ethylene***

The ground state geometry of ethylene is planar, and we use xy as the molecular plane in our calculations. At this geometry, the two lowest-lying valence excited states are the $^1B_{1u}/\pi 3s$ (S$_1$) and $^1B_{3u}/\pi\pi^*$ (S$_2$) states, with excitation energies of 7.44 eV and 8.06 eV, respectively, calculated at the EOM-CCSD/aug-cc-pVTZ level.

The ground-state experimental cross-section of ethylene has three main peaks below the ionization threshold, characterized as $1s_\mathrm{C} \rightarrow \pi^*$ at 284.7 eV, $1s_\mathrm{C} \rightarrow 3s, 3p$ at 287.8 eV, and $1s_\mathrm{C} \rightarrow 4p$ at 289.3 eV [15]. As in the previous cases, the pre-ionization threshold region of the core excitations for S$_1$ and S$_2$ states is dominated by one relatively intense peak, as seen in Fig. 8. Unlike the two previous molecules, however, the characteristic $1s \rightarrow$ SOMO signature peaks of the two valence excited states (at the Franck-Condon geometry) in ethylene are separated by more than 2 eV, even though the lower-lying SOMO in the valence excitation configuration is the same π orbital. This is because the energy difference between the two ground-state core-excited configurations that are equal to the two final configurations of the valence excited states is not cancelled out by the energy difference between the two valence excited states (according to our calculations, $1b_{3u} \rightarrow 2b_{2g} = 285.08$ eV and $1b_{3u} \rightarrow 4a_g$

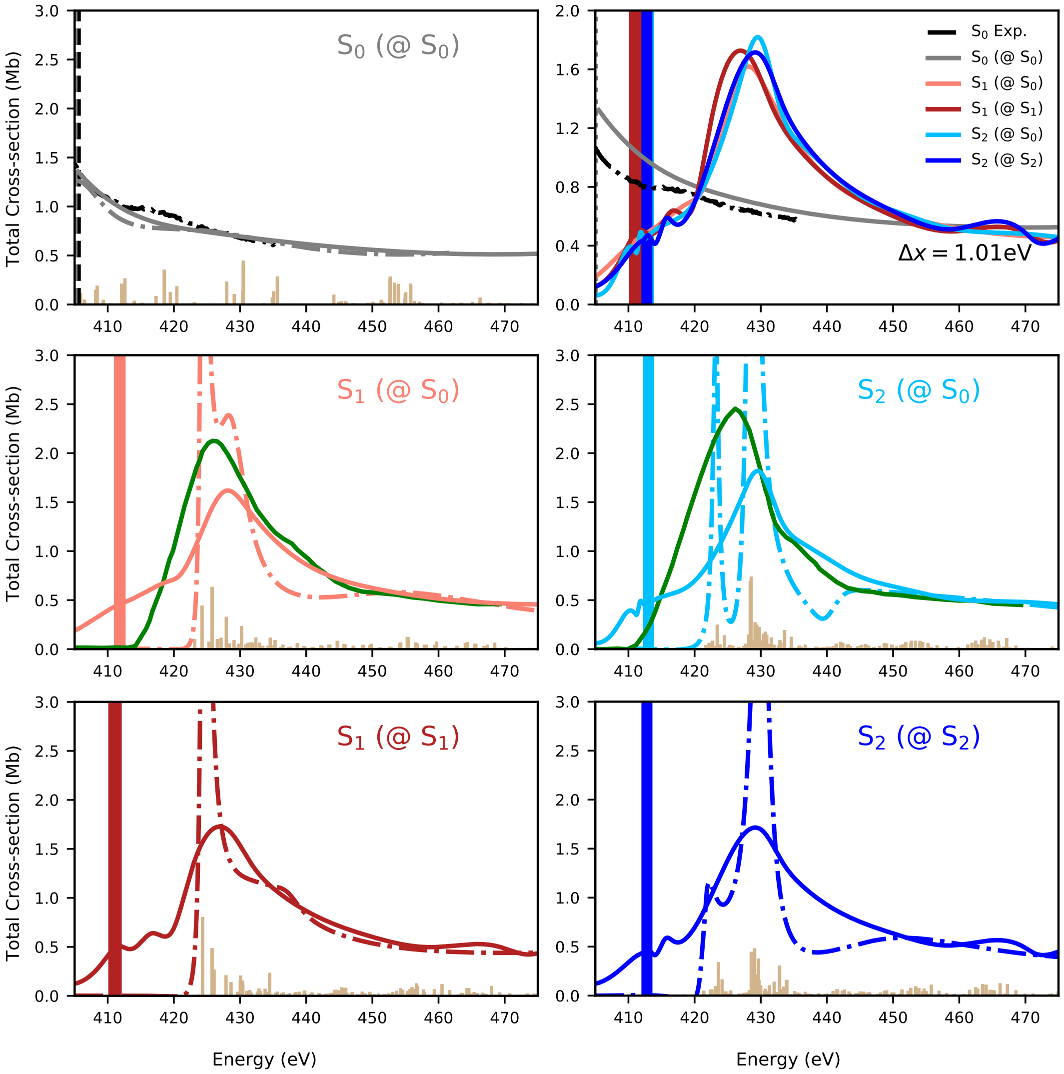

Figure 7.: NH$_3$. Ground and excited state nitrogen K-edge total cross-sections obtained with chain length $k = 500$ using EOM-CCSD/aug-cc-pVTZ+KBJ$(7s7p7d)_{n=2-8}$. For more details on color code and styles used here, see caption of Fig. 5. The computed spectra are shifted by $\Delta x = 1.01$ eV to match the position of the first ground state excitation peak (see Fig. 6). Experimental results are taken from Ref. 73.

= 287.64 eV for the ground state). The intensity of the first absorption peak lowers in value from the ground state to S$_1$ to S$_2$ state. A shoulder-like feature is found for excitation from S$_1$ state. A total of 5 peaks are seen for core-excitation from S$_2$ state. At the CVS-ADC(2)-x level [15], one dominant peak is observed in the pre-edge region for both low-lying S$_1$ and S$_2$ excited states (Fig. 2 of Ref. 15). On the other hand, CVS-DFT/MRCI [14] yields 3 peaks in the pre-edge region of excitation from the S$_1$ state at the Franck-Condon geometry. The varying number of peaks in the XAS from the valence excited state is an interesting feature of the pre-edge part, which could also possibly be used as a probe of the electronic character of the initial state.

As anticipated, in addition to the Frank-Condon geometry, we considered a "twisted-pyramidized" (Tw-py) geometric structure, see Computational Details. We labeled this

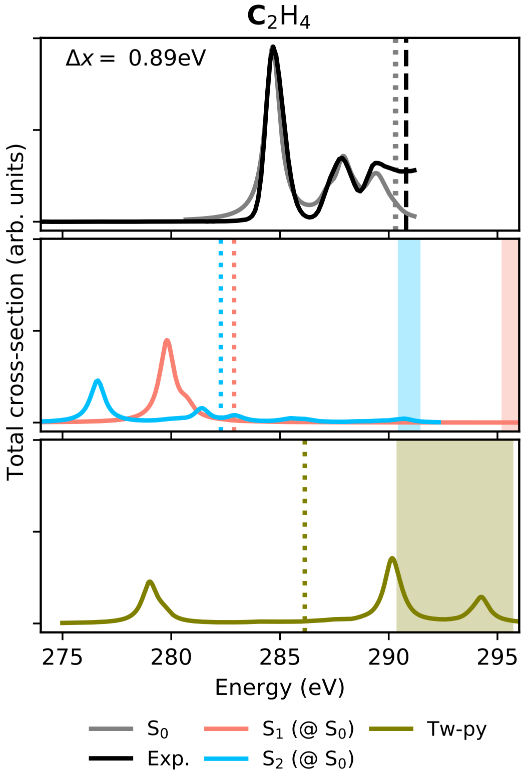

Figure 8.: C_2H_4. Ground and excited state carbon K-edge photoabsorption cross-section obtained with chain length $k = 500$ using EOM-CCSD/aug-cc-pVTZ+KBJ$(7s7p7d)_{n=2-8}$. Experimental results are taken from Ref. 73. Pale green color denotes the XA spectra obtained using the conical intersection geometry (Tw-py). For additional details on color code and styles used here, see caption of Fig. 2.

structure Tw-py, since it should be somewhat similar to the Tw-py minimal energy conical intersection geometry of Ref. 77. At this geometry, we find the lowest-energy EOM-CCSD S_1 state to be of $\pi\pi^*$ character. The (bound) XAS spectrum for this structure, see bottom panel of Fig. 8, has two relatively intense peaks roughly 10 eV apart, which is slightly larger than what was found by Neville *et al.* at CVS-ADC(2)-x level [77].

The continuum region of the valence excited states' cross-section at the FC geometry is characterized by broad, relatively intense, peaks at around 310-315 eV. As for the molecules discussed earlier, this is in contrast to the ground state XAS, for which the continuum part is flatter. Broad bands were obtained at around 305 eV with CVS-ADC(2)-x [15], once again red-shifted by approximately 10 eV with relation to our CVS-EOM-CCSD peaks. Neville *et al.* attributed the broad feature of the $S_1\pi3s$ state to shape resonances, i.e., excitations into one or more quasi-bound states. Previous computational studies have suggested that $1s \to \sigma^*_{\mathrm{C=C}}$ transitions contribute to shape resonances in the ground state XAS of ethylene [78].

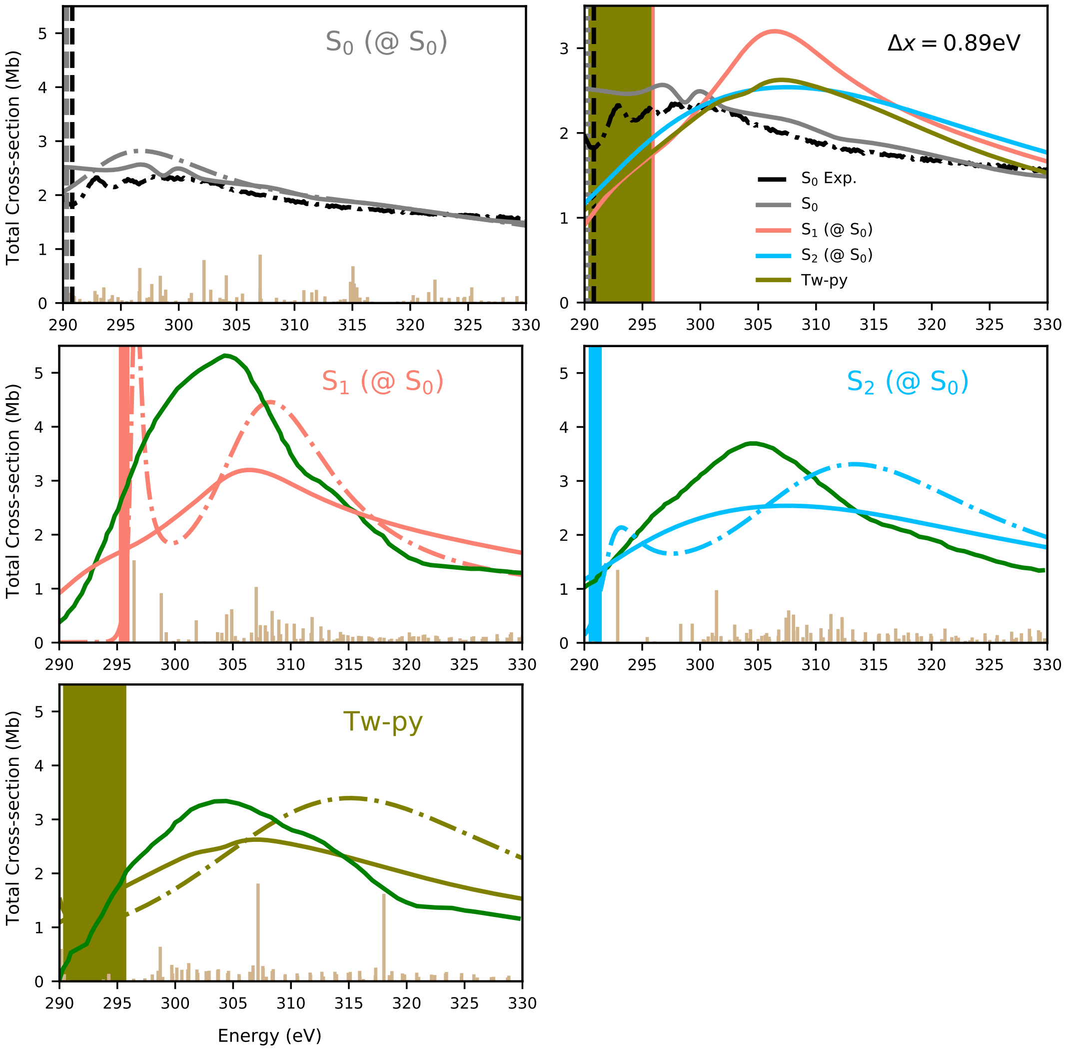

Figure 9.: C_2H_4. Ground and excited state carbon K-edge total cross-sections obtained with chain length $k = 500$ using EOM-CCSD/aug-cc-pVTZ+KBJ$(7s7p7d)_{n=2-8}$. Experimental results are taken from Ref. 73. Pale green is used to denote the spectra obtained using the Tw-py (conical intersection) geometry. For more details on color code and styles used here, see caption of Fig. 5.

4.5. Uracil

As last system, we considered the XA for both ground state and the first two valence excited states of uracil ($C_4H_4N_2O_2$) at the carbon, nitrogen and oxygen K-edges. The first valence excited state of uracil is of $n_O\pi^*$ character and optically dark, whereas the second, bright, excited state is of $\pi\pi^*$ character. A detailed analysis of the pre-edge region at the fc-CVS-EOM-CCSD level has been presented by Vidal *et al.* [11], which includes simulations of the XA spectra of the excited states S_1 and S_2. Their results are in line with our computations. Transient-state XA of uracil has also been investigated at the MOM-TDDFT and MOM-CCSD level by Tsuru *et al.* [66], and at the RASPT2 level by Mukamel and coworkers [3].

The ground-state carbon K-edge shows four well resolved $1s \rightarrow \pi^*$ excitations in the

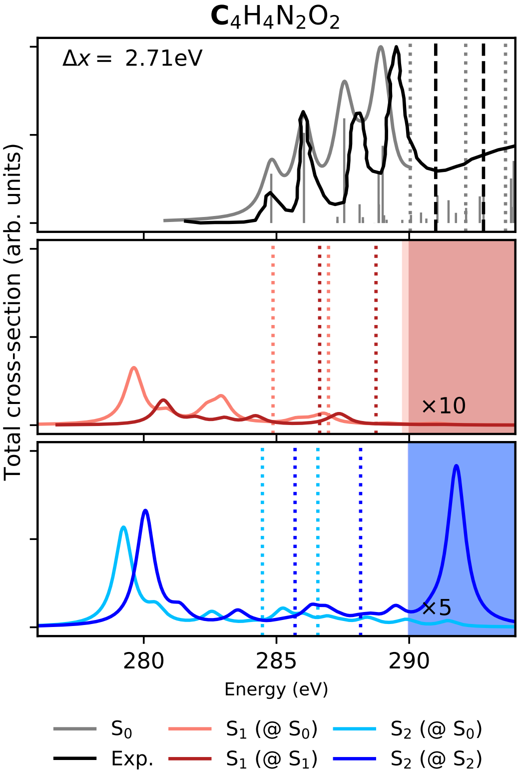

Figure 10.: $C_4H_4N_2O_2$. Carbon K-edge ground and excited state pre-edge XA cross-sections obtained with chain length k=500 using EOM-CCSD/aug-cc-pCVDZ(C)/aug-cc-pVDZ(N,O;H) + KBJ$(7s7p7d)_{n=2-8}$. Experimental results are taken from Ref. 79 and 72. Note the enhancing factors for the excited states' XAS.

pre-edge region. Experimentally they are positioned at 284.7, 285.9, 288.0 and 289.3 eV according to Ref. 79 and at 284.79, 286.11, 288.00, 289.50 eV according to Ref. 72. The four peaks are primarily due to transition from the four non-equivalent carbon atoms to π^* orbitals [72, 79]. With the somewhat limited basis set adopted here, we obtain the first transition at 287.51 eV; the second and third peaks are at 288.75 and 290.00 eV. Multiple closely packed, less intense, transitions also contribute, in particular to the fourth peak. In line with the findings of Ref. [11], at the carbon K-edge the intensity of the core excitations from the S_1 state is quite low. The photoabsorption cross-section from the S_2 state is larger and it shows one major peak at around 280 eV with a shoulder and a number of multiple peaks of lower intensity at higher energies (see Fig. 10).

The cross-sections above the ionization threshold are again rather different for ground and excited states, as shown in Figure 11. The ground state cross-section shows two broad peaks above the ionization limit, that have been attributed to σ^* shape resonances [72, 79]. In this case, the Padé approximant approach yields sharper peaks than Stieltjes Imaging, with the latter being slightly more in line with the experiment. The cross-sections above ionization for the S_1 and S_2 excited states show a broad feature (split in two in the case of the Stieltjes cross-section of S_2) peaked at almost the same energy, which is some 5-10 eV higher that the ground-state shape resonances. This makes it difficult to distinguish between the two valence excited states

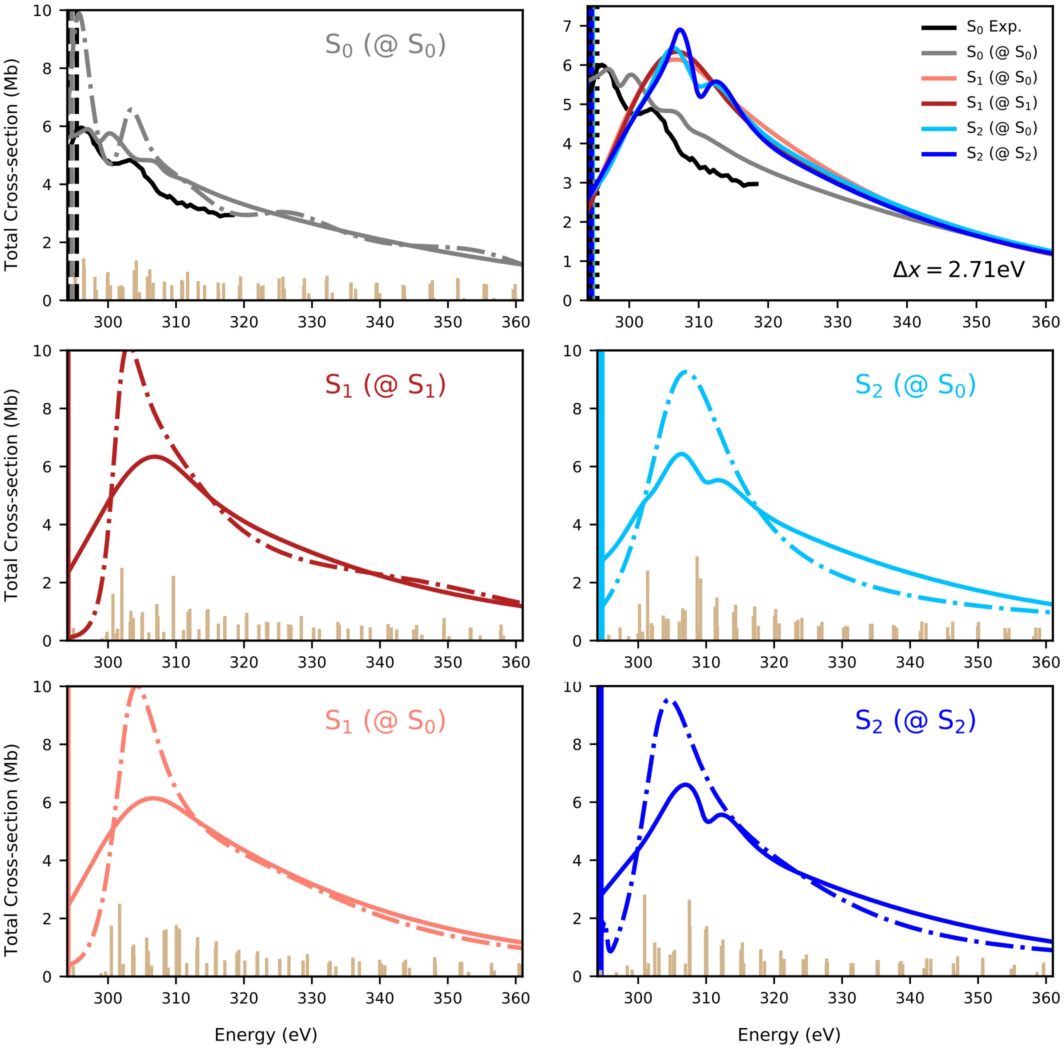

Figure 11.: Uracil. Carbon K-edge ground and excited state X-ray photo-ionization cross-sections obtained with chain length $k = 500$ using EOM-CCSD/aug-cc-pCVDZ(C)/aug-cc-pVDZ(N,O,H) + KBJ$(7s7p7d)_{n=2-8}$. Experimental results are taken from Ref. 79 and 72.

solely based on their photoionization cross-sections at the carbon K-edge.

The nitrogen K-edge photoabsorption spectra of uracil are shown in Figure 12. It can be subdivided into 2 regions of π^* resonances at 398-404 eV and σ^* resonances at 405, 407-408 and 416 eV, in the experimental cross-sections [80]. The core excitations to π^* orbitals in the ground state are experimentally observed at 401.0 and 402.1 eV [79]. The computed spectrum is here shifted by 5.4 eV and is an envelope over multiple excitation energies and oscillator strengths. The first two σ^* transitions are known to be due to $1s_\mathrm{N} \rightarrow \sigma^*$(N-C) transitions and the last one is a characteristic feature of molecules containing amino groups, previously assigned to $1s_\mathrm{N} \rightarrow \sigma^*$(N-C=O) transitions[81, 82]. The second and third σ^* transitions lie above the ionization threshold and are also a characteristic feature of core excitations from the ground state. The pre-edge resonance regions of the S$_1$ and S$_2$ states have one dominant peak and

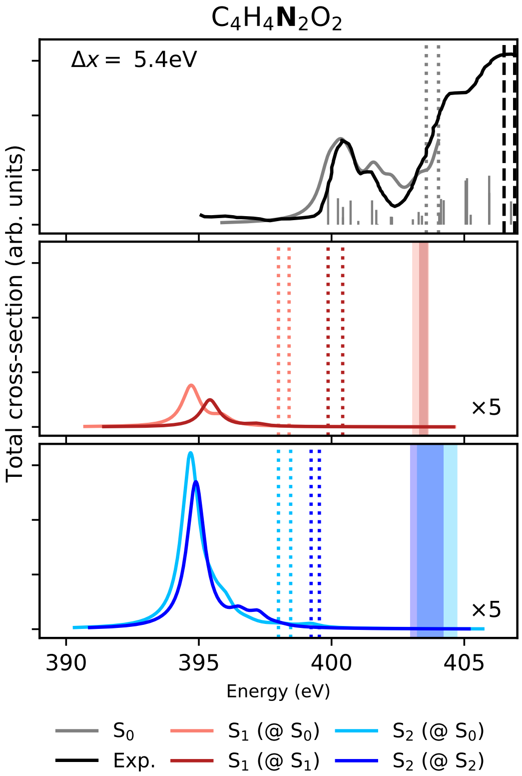

Figure 12.: Uracil: Nitrogen K-edge ground and excited state pre-edge XA cross-sections obtained with chain length $k = 500$ using EOM-CCSD/aug-cc-pCVDZ(N)/aug-cc-pVDZ(C,H,O) + KBJ$(7s7p7d)_{n=2-8}$. Experimental results are taken from Ref. 79, 80 and 72. Note the enhancing factor for the excited states' XAS.

a shoulder. The S_2 peak has clearly higher intensity, due to a more favorable overlap of the $1s_\mathrm{N}$ hole with the π particle orbital than with the n one, see also Ref. [11].

Turning to the region above the ionization limit, the σ^* resonances clearly yield a split feature at around 410 eV in the (shifted) ground-state cross-section. This is not captured by the quadrature procedures, but well reproduced by Lorentzian broadening of the stick spectrum just above the ionization limit. As for the excited states, a rather intense band is obtained around 420 eV for both. Some shoulder-like structure is seen to emerge around 410 eV, especially in the Stieltjes case. Additionally, the Padé approach generates a broad feature at around 450 eV for the ground state, and at 460-470 eV for the excited states.

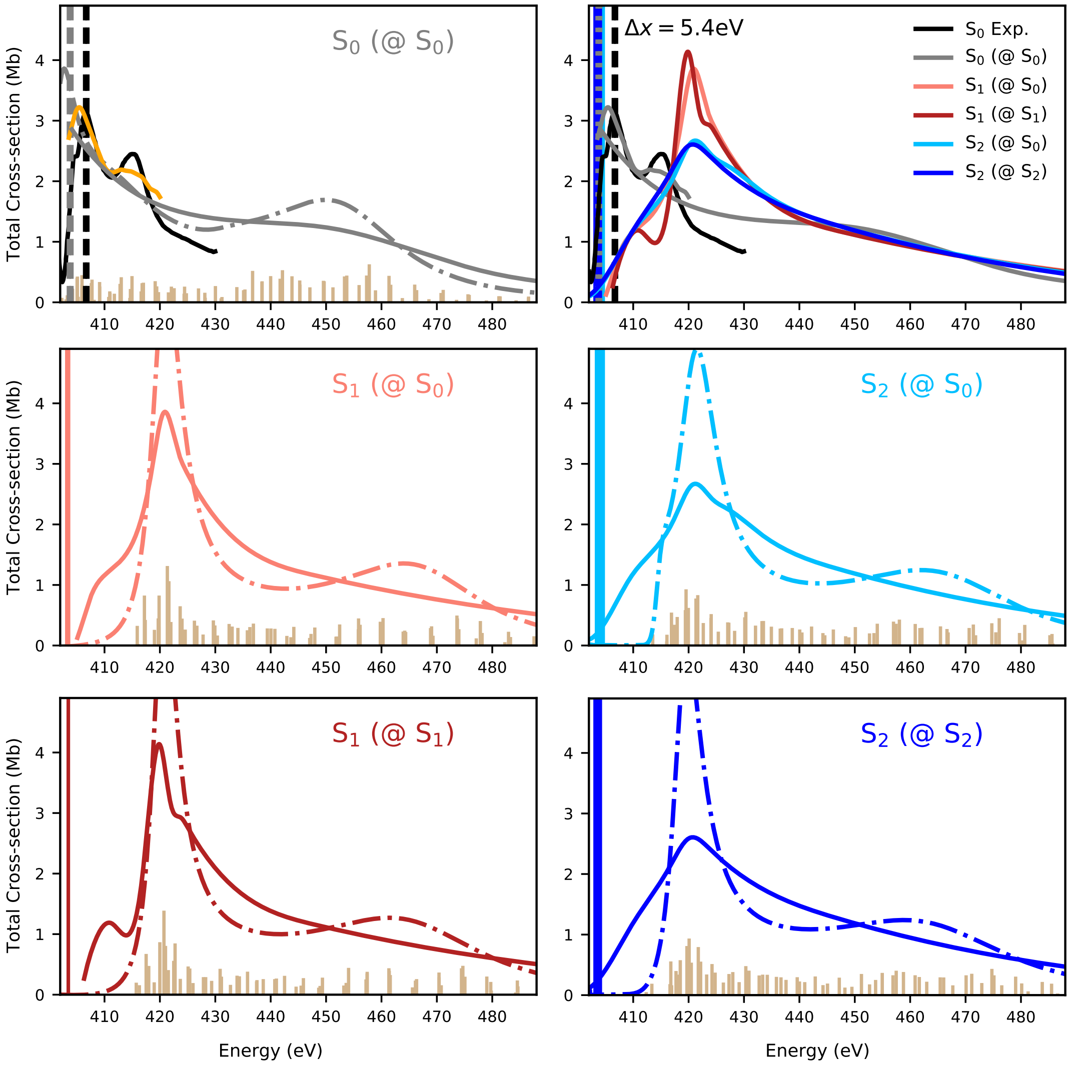

Figure 13.: Uracil: Nitrogen K-edge ground and excited state X-ray photo-ionization cross-sections obtained with chain length $k = 500$ using EOM-CCSD/aug-cc-pCVDZ(N)/aug-cc-pVDZ(C,O,H) + KBJ$(7s7p7d)_{n=2-8}$. Experimental results are taken from Ref. 79 and 72. The cross-section in orange in the top-left panel is from the Lorentzian broadening of spectral sticks close to the ionization limit.

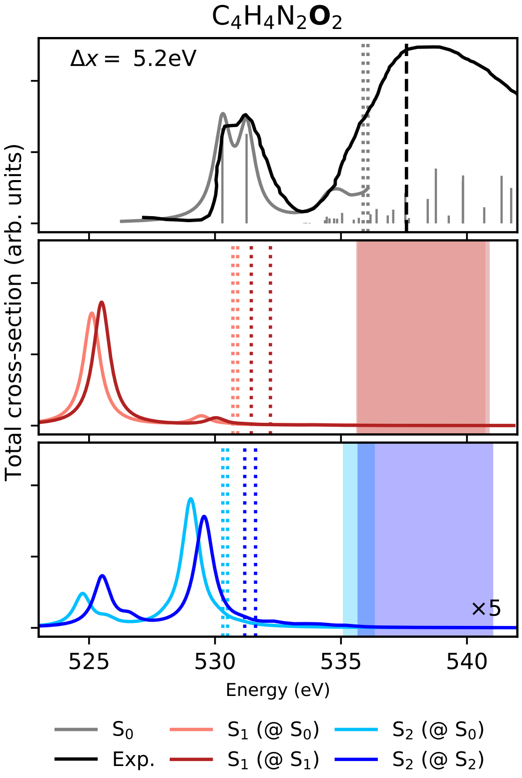

Figure 14.: Uracil: Oxygen K-edge ground and excited state X-ray photo-absorption cross-sections below the ionization limit, obtained with chain length $k = 500$ using EOM-CCSD/aug-cc-pCVDZ(O)/aug-cc-pVDZ(C,H,N) + KBJ$(7s7p7d)_{n=2-8}$. Experimental results are taken from Ref. 79, 80 and 71. Note the enhancing factor for the S_2 XAS.

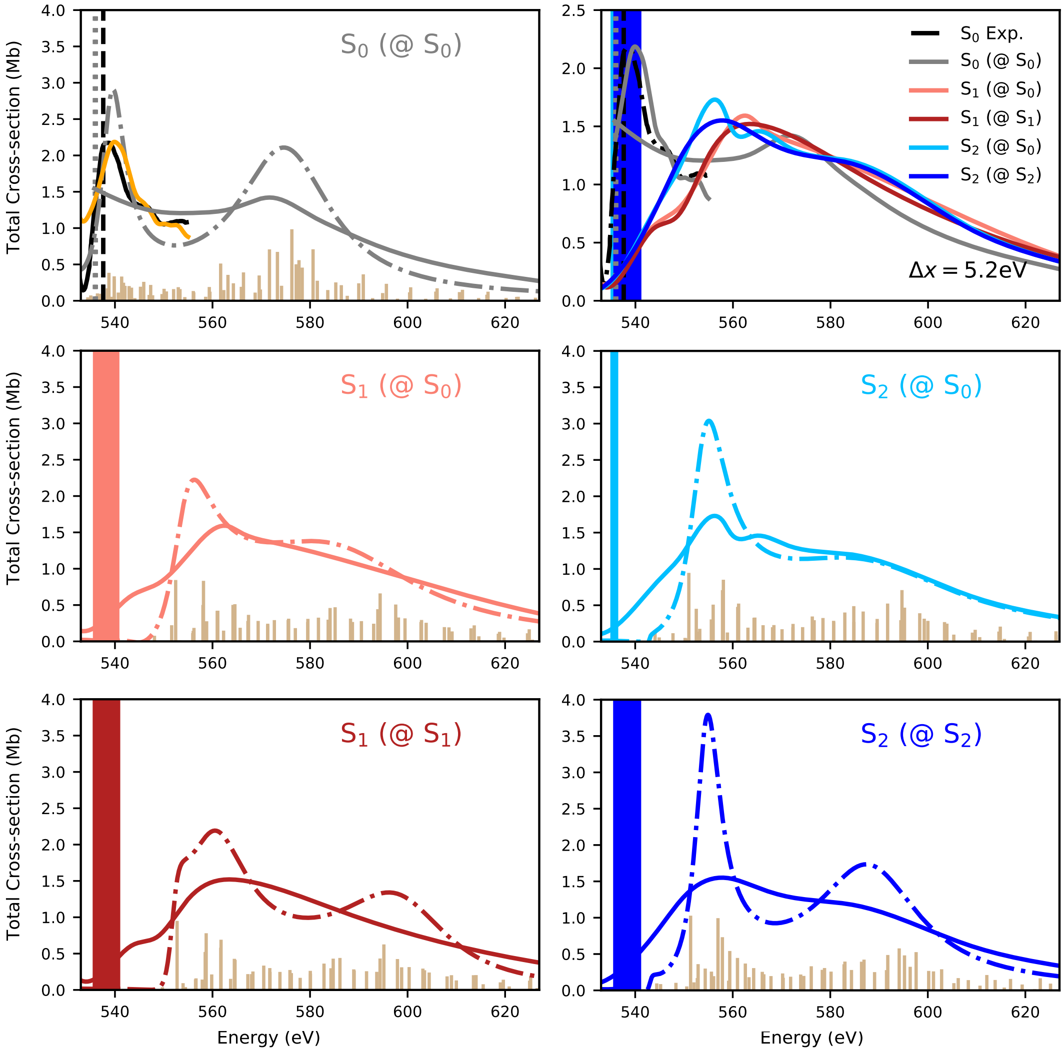

Figure 15.: Uracil. Oxygen K-edge ground and excited state X-ray photoionization cross-sections obtained with chain length k=500 using EOM-CCSD/aug-cc-pCVDZ(O)/aug-cc-pVDZ(C,N,H) + KBJ$(7s7p7d)_{n=2-8}$. Experimental results are taken from Ref. 80 and 71. The cross-section in orange is a Lorentzian broadening of the sticks close to the ionization threshold.

The ground state oxygen K-edge XA spectrum shows, like at the nitrogen K-edge, two distinct regions of lower-energy π^* (see Figure 14) and near-ionization-threshold σ^* core excitations (see Figure 15). Uracil has only carbonyl type oxygen atoms making the spectra simpler in comparison to XAS of nitrogen and carbon edges. The π^* region comprises of two closely spaced spectral peaks. This splitting is relatively small and due to the slight difference in the chemical environment of the two oxygen centers [80]. Note that the computed spectra at the oxygen K-edge are here shifted by 5.2 eV. Core excitations from the low-lying excited states result in two peaks. Absorption from S$_1$ shows an intense peak at 530 eV (525 eV in the shifted spectrum). The S$_2$ XAS has two peaks, a weaker one overlapping with the intense peak of S$_1$ and a second, more intense one, at 535 eV ($\approx$ 530 eV in the shifted spectrum). The very intense S$_1$ band (due to the high-locality of the $1s_{\mathrm{O}} \rightarrow n_{\mathrm{O}}$ transition) could be exploited to experimentally detect the population of S_1 state in the time-resolved XAS spectra of uracil, as done in the experimentally investigated case of thymine [4, 83].

The dominant and broad σ^* components for ground state ionization appear at 539-540 eV [79]. As for the previous edge, this is not captured by the Stieltjes cross-section, but well reproduced by simple Lorentzian broadening. Stieltjes Imaging generates a broad peak at around 560 eV for the excited state molecule. The Padé approach using optimized excited state geometries shows an additional broad peak at 590-600 eV in comparison to Franck-Condon geometry.

5. Conclusion

Previous studies [16, 17, 29, 30] have shown that quadrature protocols based on EOM-CC Lanczos pseudo-spectra can be reliably applied to compute ground-state total photoionization cross-sections in both UV and X-ray spectral regions, as well as transient excited state cross-sections in the VUV region. Here, we have presented an extension of these formulations to the inner-shell cross-sections of valence excited states.

As demonstrated in several studies, see e.g. Refs. [1–10, 13, 15, 77], the cross-sections below the ionization limit can be used as a probe of excited state dynamics. In all cases considered here, there is a substantial difference between the XAS spectra of the ground and the excited states. However, for water and ammonia molecule, the two lowest lying excited states have their 'signature' excitation features at the same energy, due to the nature of the underlying transitions. On the contrary, this is not the case for ethylene, which can be mainly attributed to the different energy of the higher energy SOMO of the valence excited configurations.

Above the ionization limit, the situation is less clear. Even though broad-band features are obtained in the photoionization cross-sections from valence excited states, it is difficult to distinguish or characterise the nature of the excited state from which they originate. The photoionization cross-sections of the different excited states here considered are mostly overlapping. This means that in an hypothetical pump-probe experiment, analysis of the inner-shell total photoionization cross-sections of the valence excited states can hardly be used to distinguish (and dynamically probe) different excited states.

Also, it is noted that where the two quadrature methods give reasonably similar results for the ground state core-ionization cross sections – and also for excited states valence ionisation, see Ref. [17] – they can produce rather different profiles for core ionization of valence excited-states. The Padé cross sections, in particular, often show sharp features, especially in correspondence to clustered, relatively intense, sticks in the underlying discretized representation of the continuum. This seems somewhat unphysical, when compared to the much smoother results of Stieltjes Imaging. Unfortunately, no experimental results are available to possibly guide in the choice of which method is most reliable. On the one hand, the Stieltjes Imaging profile depends on the number $2n$ of spectral moments used to generate the $(n-1)$ principal pseudo-spectra. Only the converged spectral moments should be employed. In practice, only a limited number is reached, after which scattered values may begin to appear. On the other hand, the Padé procedure depends on the selection of interpolating points in the complex plane, for which no unique recipe is available [16, 32, 35].

Using the S_2 state of NH_3 as a test case, we show in Figures S8 and S9 the effect of alternative selections of the order in Stjelties Imaging and of the interpolation points in the Padé based procedure. Alternative choices lead to slight different cross sections for each method, but the main spectral features remain, and in particular the presence of the two sharp peaks in the Padé-based cross-section, whose origin remains unclear to us. We therefore do not recommend the use of the Padé procedure to compute the core photoionization cross-section of valence excited states.

Finally, the moment theory here applied to describe the ionization cross-section implies that bound and continuum parts are separable by the ionization threshold. Where this threshold should be placed in the case of core ionization from a valence excited state remains somewhat unclear. The ionization energy for the shake-down transition by which the molecule returns on the ground-state core-ionized configuration is straightforwardly found within the vertical approximation (Δ values), which thus gives the excited-state threshold at lower energy than the one of the ground state. The threshold for the ($2h$-$1p$) configuration corresponding to the most intensity is at higher energy. To overcome the difficulty of regular EOM-CCSD to yield accurate values for this 'doubly excited' configuration, we here explored different flavours of a two-step, MOM-based approach to estimate a 'threshold window'. We feel this requires further detailed investigation.

Disclosure statement

The authors declare no conflict of interest.

Acknowledgments

We acknowledge useful discussions with Prof. Piero Decleva (University of Trieste) and Dr. Michael Schuurman (National Research Council of Canada / University of Ottawa).

Funding

This project has received funding from the European Unions Horizon 2020 research and innovation programme under the Marie Skodowska-Curie grant agreement No 765739, "COSINE–European Training Network on COmputational Spectroscopy In Natural sciences and Engineering". S.C. and B.N.C.T acknowledge financial support from the Independent Research Fund Denmark–Natural Sciences, DFF-RP2 Grant No. 7014-00258B.

References

[1] A. Bhattacherjee, K. Schnorr, S. Oesterling, Z. Yang, T. Xue, R. de Vivie-Riedle and S.R. Leone, J. Am. Chem. Soc. **140** (39), 12538–12544 (2018).

[2] H.R. Hudock, B.G. Levine, A.L. Thompson, H. Satzger, D. Townsend, N. Gador, S. Ullrich, A. Stolow and T.J. Martínez, J. Phys. Chem. A **111** (34), 8500–8508 (2007).

[3] W. Hua, S. Mukamel and Y. Luo, J. Phys. Chem. Lett. **10** (22), 7172–7178 (2019).

[4] T.J.A. Wolf, R.H. Myhre, J.P. Cryan, S. Coriani, R.J. Squibb, A. Battistoni, N. Berrah, C. Bostedt, P. Bucksbaum, G. Coslovich, R. Feifel, K.J. Gaffney, J. Grilj, T.J. Martinez, S. Miyabe, S.P. Moeller, M. Mucke, A. Natan, R. Obaid, T. Osipov, O. Plekan, S. Wang, H. Koch and M. Gühr, Nature Commun. **8** (29), 1–7 (2017).

[5] G.C. ONeil, L. Miaja-Avila, Y.I. Joe, B.K. Alpert, M. Balasubramanian, D.M. Sagar, W. Doriese, J.W. Fowler, W.K. Fullagar, N. Chen, G.C. Hilton, R. Jimenez, B. Ravel, C.D. Reintsema, D.R. Schmidt, K.L. Silverman, D.S. Swetz, J. Uhlig and J.N. Ullom, J. Phys. Chem. Lett. **8** (5), 1099–1104 (2017).

[6] G. Olivo, A. Barbieri, V. Dantignana, F. Sessa, V. Migliorati, M. Monte, S. Pascarelli, T. Narayanan, O. Lanzalunga, S. Di Stefano and P. DAngelo, J. Phys. Chem. Lett. **8** (13), 2958–2963 (2017).

[7] G. Subramanian, X. Zhang, G. Kodis, Q. Kong, C. Liu, A. Chizmeshya, U. Weierstall and J. Spence, J. Phys. Chem. Lett. **9** (7), 1542–1546 (2018).

[8] A.D. Smith, T. Baliunas, Y.P. Chang, C. Schmidt, K. Zinchenko, F.B. Nunes, E. Rossi, V. Svoboda, Z. Yin, J.P. Wolf and H.J. Wörner, J. Phys. Chem. Lett. **11** (6), 1981–1988 (2020).

[9] C. Kleine, M. Ekimova, G. Goldsztejn, S. Raabe, C. Strüber, J. Ludwig, S. Yarlagadda, S. Eisebitt, M.J.J. Vrakking, T. Elsaesser, E.T.J. Nibbering and A. Rouze, J. Phys. Chem. Lett. **10** (1), 52–58 (2019).

[10] B.E. Van Kuiken, H. Cho, K. Hong, M. Khalil, R.W. Schoenlein, T.K. Kim and N. Huse, J. Phys. Chem. Lett. **7** (3), 465–470 (2016).

[11] M.L. Vidal, X. Feng, E. Epifanovsky, A. Krylov and S. Coriani, J. Chem. Theory Comput. **15**, 3117–3133 (2019).

[12] S. Tsuru, M.L. Vidal, M. Pápai, A.I. Krylov, K.B. Møller and S. Coriani, J. Chem. Phys. **151** (12), 124114 (2019).

[13] N.H. List, A.L. Dempwolff, A. Dreuw, P. Norman and T.J. Martínez, Chem. Sci. **11**, 4180–4193 (2020).

[14] I. Seidu, S.P. Neville, M. Kleinschmidt, A. Heil, C.M. Marian and M.S. Schuurman, J. Chem. Phys. **151** (14), 144104 (2019).

[15] S.P. Neville, V. Averbukh, M. Ruberti, R. Yun, S. Patchkovskii, M. Chergui, A. Stolow and M.S. Schuurman, J. Chem. Phys. **145** (14), 144307 (2016).

[16] B.N.C. Tenorio, T. Moitra, M.A.C. Nascimento, A.B. Rocha and S. Coriani, J. Chem. Phys. **150** (22), 224104 (2019).

[17] B.N.C. Tenorio, M.A.C. Nascimento, A.B. Rocha and S. Coriani, J. Chem. Phys. **151** (18), 184106 (2019).

[18] P. Langhoff, Chem. Phys. Lett. **22** (1), 60–64 (1973).

[19] P. Langhoff, C. Corcoran, J. Sims, F. Weinhold and R. Glover, Phys. Rev. A **14** (3), 1042–1056 (1976).

[20] V. Carravetta, H. Ågren, H.J.Å. Jensen, P. Jørgensen and J. Olsen, J. Phys. B: At. Mol. Opt. Phys. **22**, 2133–2140 (1989).

[21] M.A.C. Nascimento, J. Mol. Struct. **120**, 227–240 (1985).

[22] B.N.C. Tenorio, S. Coriani, A.B. Rocha and M.A.C. Nascimento, in *Advances in Methods and Applications of Quantum Systems in Chemistry, Physics, and Biology*, edited by Alexander V. Glushkov, Olga Yu. Khetselius, Jean Maruani and Erkki Brändas (Springer International Publishing, Cham, 2021), pp. 151–179.

[23] H. Bachau, E. Cormier, P. Decleva, J.E. Hansen and F. Martín, Rep.Prog. Phys. **64** (12), 1815–1943 (2001).

[24] T. Moitra, A. Ponzi, H. Koch, S. Coriani and P. Decleva, J. Phys. Chem. Lett. **11** (13), 5330–5337 (2020).

[25] T. Moitra, S. Coriani and P. Decleva, J. Chem. Theory Comput. (2021).

[26] K. Gokhberg, V. Vysotskiy, L.S. Cederbaum, L. Storchi, F. Tarantelli and V. Averbukh, J. Chem. Phys. **130**, 064104 (2009).

[27] S. Kopelke, K. Gokhberg, L.S. Cederbaum, F. Tarantelli and V. Averbukh, J. Chem. Phys. **134** (2), 024106 (2011).

[28] S. Kopelke, K. Gokhberg, V. Averbukh, F. Tarantelli and L.S. Cederbaum, J. Chem. Phys. **134**, 094107 (2011).

[29] J. Cukras, S. Coriani, P. Decleva, O. Christiansen and P. Norman, J. Chem. Phys. **139**, 094103 (2013).

[30] J. Cukras, P. Decleva and S. Coriani, J. Chem. Phys. **141**, 174315 (2014).

[31] B.N.C. Tenorio, M.A.C. Nascimento, S. Coriani and A.B. Rocha, J. Chem. Theory Comp. **12** (9), 4440–4459 (2016).

[32] B.N.C. Tenorio, M.A.C. Nascimento and A.B. Rocha, J. Chem. Phys. **148** (7), 074104 (2018).

[33] B.N. Cabral Tenorio, M.A. Chaer Nascimento and A.B. Rocha, J. Phys. Chem. A. **124** (13), 2591–2600 (2020).

[34] B.N.C. Tenorio, M.A.C. Nascimento and A.B. Rocha, J. Chem. Phys. **150** (15), 154308 (2019).

[35] B.N.C. Tenorio, R.R. Oliveira, M.A.C. Nascimento and A.B. Rocha, J. Chem. Theory Comput. **14** (10), 5324–5338 (2018).

[36] M. Ruberti, R. Yun, K. Gokhberg, S. Kopelke, L.S. Cederbaum, F. Tarantelli and V. Averbukh, J. Chem. Phys. **140** (18), 184107 (2014).

[37] M. Ruberti, R. Yun, K. Gokhberg, S. Kopelke, L.S. Cederbaum, F. Tarantelli and V. Averbukh, J. Chem. Phys. **139**, 144107 (2013).

[38] J.F. Stanton and R.J. Bartlett, J. Chem. Phys. **98** (9), 7029–7039 (1993).

[39] O. Christiansen, P. Jørgensen and C. Hättig, Int. J. Quantum Chem. **98**, 1 (1998).

[40] H. Koch and P. Jørgensen, J. Chem. Phys. **93**, 3333–3344 (1990).

[41] F. Pawłowski, J. Olsen and P. Jørgensen, J. Chem. Phys. **142** (11), 114109 (2015).

[42] S. Coriani, F. Pawłowski, J. Olsen and P. Jørgensen, J. Chem. Phys. **144** (2), 024102 (2016).

[43] R. Faber and S. Coriani, J. Chem. Theory Comput. **15** (1), 520–528 (2019).

[44] A.I. Krylov, Ann. Rev. Phys. Chem. **59** (1), 433–462 (2008).

[45] T. Helgaker, P. Jørgensen and J. Olsen, *Molecular Electronic-Structure Theory* (Wiley, Chichester, 2004).

[46] S. Coriani, T. Fransson, O. Christiansen and P. Norman, J. Chem. Theory Comp. **8**, 1616–1628 (2012).

[47] S. Coriani, O. Christiansen, T. Fransson and P. Norman, Phys. Rev. A **85**, 022507 (2012).

[48] G.H. Golub and C.F. Van Loan, *Matrix Computations*, 3rd ed. (Johns Hopkins University Press, Philadelphia, PA, 1996).

[49] L.S. Cederbaum, W. Domcke and J. Schirmer, Phys. Rev. A: At. Mol. Opt. Phys. **22**, 206–222 (1980).

[50] S. Coriani and H. Koch, J. Chem. Phys. **143** (18), 181103 (2015).

[51] F. Müller-Plathe and G.H. Diercksen, in *Electronic Structure of Atoms, Molecules and Solids. Proceeding of the II Escola Brasileira de Estructure Eletronica, Olinda, Brazil, July 17-22, 1989*, edited by Sylvio Canuto, Jose D'Albuquerque e Castro and Fernando J. Paixao (World Scientific, Olinda, Brazil, 1990), pp. 1–29.

[52] K. Aidas, C. Angeli, K.L. Bak, V. Bakken, R. Bast, L. Boman, O. Christiansen, R. Cimiraglia, S. Coriani, P. Dahle, E. 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. Hjertenaes, S. Høst, I.M. Høyvik, M.F. Iozzi, B. Jansík, H. Jensen, D. Jonsson, P. Jørgensen, J. Kauczor, S. Kirpekar, T. Kjaergaard, W. Klopper, S. Knecht, R. Kobayashi, H. Koch, J. Kongsted, A. Krapp, K. Kristensen, A. Ligabue, O. Lutnæs, J. Melo, K. Mikkelsen, R. Myhre, C. Neiss, C. Nielsen, P. Norman, J. Olsen, J. Olsen, A. Osted, M. Packer, F. Pawlowski, T. Pedersen, P. Provasi, S. Reine, Z. Rinkevicius, T.A. Ruden, K. Ruud, V.V. Rybkin, P. Sałek, C.C.M. Samson, A.S. de Merás, T. Saue, S. 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, WIREs Comput. Mol. Sci. **4**, 269–284 (2014).

[53] J.M.H. Olsen, S. Reine, O. Vahtras, E. Kjellgren, P. Reinholdt, K.O. Hjorth Dundas, X. Li, J. Cukras, M. Ringholm, E.D. Hedegård, R. Di Remigio, N.H. List, R. Faber, B.N. Cabral Tenorio, R. Bast, T.B. Pedersen, Z. Rinkevicius, S.P.A. Sauer, K.V. Mikkelsen, J. Kongsted, S. Coriani, K. Ruud, T. Helgaker, H.J.A. Jensen and P. Norman, J. Chem.

Phys. **152** (21), 214115 (2020).

[54] Dalton, a molecular electronic structure program, Release Dalton2020.alpha (2020) See http://daltonprogram.org.

[55] D.A. Matthews, L. Cheng, M.E. Harding, F. Lipparini, S. Stopkowicz, T.C. Jagau, P.G. Szalay, J. Gauss and J.F. Stanton, J. Chem. Phys. **152** (21), 214108 (2020).

[56] Y. Shao, Z. Gan, E. Epifanovsky *et al.*, Mol. Phys. **113** (2), 184–215 (2015).

[57] K.B. Wiberg, Y.g. Wang, A.E. de Oliveira, S.A. Perera and P.H. Vaccaro, J. Phys. Chem. A **109** (3), 466–477 (2005).

[58] T.H. Dunning, J. Chem. Phys. **90**, 1007 (1989).

[59] D.E. Woon and T.H. Dunning, J. Chem. Phys. **103** (11), 4572–4585 (1995).

[60] R.A. Kendall, T.H. Dunning and R.J. Harrison, J. Chem. Phys. **96** (9), 6796–6806 (1992).

[61] K. Kaufmann, W. Baumeister and M. Jungen, J. Phys. B: At. Mol. Opt. Phys. **22**, 2223–2240 (1989).

[62] J. Stanton and G. J., J. Chem. Phys. **111**, 8785 (1999).

[63] S. Coriani and H. Koch, J. Chem. Phys. **145** (14), 149901 (2016).

[64] A.T.B. Gilbert, N.A. Besley and P.M.W. Gill, J. Phys. Chem. A **112** (50), 13164–13171 (2008).

[65] G.M.J. Barca, A.T.B. Gilbert and P.M.W. Gill, J. Chem. Theory Comput. **14** (3), 1501–1509 (2018).

[66] S. Tsuru, M.L. Vidal, M. Pápai, A.I. Krylov, K.B. Møller and S. Coriani, Structural Dynamics **8** (2), 024101 (2021).

[67] M.L. Vidal, A.I. Krylov and S. Coriani, Phys. Chem. Chem. Phys. **22**, 2693–2703 (2020).

[68] M.L. Vidal, A.I. Krylov and S. Coriani, Phys. Chem. Chem. Phys. **22**, 3744–3747 (2020).

[69] W. Jolly, K. Bomben and C. Eyermann, At. Data and Nucl. Data Tables **31** (3), 433–493 (1984).

[70] K. Siegbahn, C. Nordling, G. Johansson, J. Hedman, P.F. Hedén, K. Hamrin, U. Gelius, T. Bergmark, L.O. Werme and Y. Baer, *ESCA Applied to Free Molecules* (North Holland Publishing Co., Amsterdam, 1969).

[71] V. Feyer, O. Plekan, R. Richter, M. Coreno, G. Vall-llosera, K.C. Prince, A.B. Trofimov, I.L. Zaytseva, T.E. Moskovskaya, E.V. Gromov and J. Schirmer, J. Phys. Chem. A **113** (19), 5736–5742 (2009).

[72] V. Feyer, O. Plekan, R. Richter, M. Coreno, M. de Simone, K.C. Prince, A.B. Trofimov, I.L. Zaytseva and J. Schirmer, J. Phys. Chem. A **114** (37), 10270–10276 (2010).

[73] A. Hitchcock and D. Mancini, J. Electron Spectrosc. **67**, 1 (1994).

[74] J. Pálenková, M. Kraus, P. Neogrády, V. Kellö and M. Urban, Mol. Phys. **106** (20), 2333–2344 (2008).

[75] J.F. Stanton and N.S. Kadagathur, J. Chem. Phys. **102** (2), 1096–1097 (1995).

[76] J.P. Carbone, L. Cheng, R.H. Myhre, D. Matthiews, H. Koch and S. Coriani, Adv. Quant. Chem. **79**, 241 (2019).

[77] S.P. Neville, V. Averbukh, S. Patchkovskii, M. Ruberti, R. Yun, M. Chergui, A. Stolow and M.S. Schuurman, Faraday Discuss. **194**, 117–145 (2016).

[78] R.E. Farren, J. Sheehy and P. Langhoff, Chem. Phys. Lett. **177** (3), 307–314 (1991).

[79] Y. Zubavichus, A. Shaporenko, V. Korolkov, M. Grunze and M. Zharnikov, J. Phys. Chem. B **112** (44), 13711–13716 (2008).

[80] K. Fujii, K. Akamatsu, Y. Muramatsu and A. Yokoya, Nucl. Instrum. Methods Phys. Res., B **199**, 249–254 (2003).

[81] Y. Zubavichus, A. Shaporenko, M. Grunze and M. Zharnikov, J. Phys. Chem. A **109** (32), 6998–7000 (2005).

[82] M.L. Gordon, G. Cooper, C. Morin, T. Araki, C.C. Turci, K. Kaznatcheev and A.P. Hitchcock, J. Phys. Chem. A **107** (32), 6144–6159 (2003).

[83] T.J.A. Wolf, A.C. Paul, S.D. Folkestad, R.H. Myhre, J.P. Cryan, N. Berrah, P.H. Bucksbaum, S. Coriani, G. Coslovich, R. Feifel, T.J. Martinez, S.P. Moeller, M. Mucke, R. Obaid, O. Plekan, R.J. Squibb, H. Koch and M. Gühr, Faraday Discuss. **228**, 555–570 (2021).