ASTRA: Transition-density-matrix approach to molecular ionization

Juan M. Randazzo [1,2] Carlos Marante [1] Siddhartha Chattopadhyay [1] Barry I. Schneider [3]
Jeppe Olsen,[4] and Luca Argenti [1,*]

[1]*Department of Physics & CREOL, University of Central Florida, Orlando, Florida 32816, USA*
[2]*Centro Atómico Bariloche, CNEA, CONICET, and Instituto Balseiro, UNCuyo, 8400 Bariloche, Argentina*
[3]*Applied and Computational Math Division, NIST, 100 Bureau Drive, Gaithersburg, Maryland 20899, USA*
[4]*Department of Chemistry, Aarhus University, Langelandsgade 140, DK 8000 Aarhus C, Denmark*

We describe ASTRA (attosecond transitions), a close-coupling approach to molecular ionization that uses many-body transition-density matrices between ionic states with arbitrary spin and symmetry, in combination with hybrid integrals between Gaussian and numerical orbitals, to efficiently evaluate photoionization observables. Within the transition-density-matrix approach, the evaluation of interchannel coupling is exact and does not depend on the size of the configuration-interaction space of the ions. Thanks to these two crucial features, ASTRA opens the way to studying highly correlated and comparatively large targets at a manageable computational cost. Here, ASTRA is used to predict the parameters of bound and autoionizing states of the boron atom and of the N_2 molecule, as well as the total photoionization cross section of boron, the nitrogen molecule (N_2), and formaldehyde (H_2CO). Our results are in excellent agreement with theoretical and experimental values from the literature.

I. INTRODUCTION

Excitation of correlated electronic motion is at the core of light-induced chemical transformations in matter [1]. The advent of attosecond spectroscopies [2–4] has made it possible to study electron dynamics on its natural sub-femtosecond timescale [5–8]. Continuous advances in the XUV and soft-x-ray ultrafast technologies, pursued at large free-electron-laser facilities [9,10] as well as in several attosecond laboratories around the world [11–16], have continued extending the range of attosecond-pulse parameters to still shorter durations, larger intensities, energies, and repetition rates [17–20]. At the same time, the focus of attosecond experiments has moved from atomic systems and small molecules to larger systems with chemical and biological relevance [21–23].

Attosecond pump-probe observables, such as coincidence photoelectron distributions and transient absorption spectra, are complex functions of the laser parameters that reflect the correlated and entangled dynamics of the probed system only indirectly. Due this complexity, *ab initio* methods beyond the single-particle approximation are essential to reconstruct and ultimately control the many-body motion excited by short light pulses [24].

Here we present ASTRA (attosecond transitions), a wave-function close-coupling approach based on a transition-density-matrix close-coupling (TDMCC) formalism, suited to describe *ab initio* the ionization processes that are made accessible by the use of new sources of coherent ultrashort light. ASTRA is designed to describe both atoms and larger molecules, for both low- and high-energy photoelectrons. It takes exchange and correlation into account to the full extent allowed by state-of-the-art large-scale configuration-interaction quantum-chemistry methods [25]. ASTRA admits a natural extension to double-ionization processes, which are needed to observe the motion of correlated electron pairs in real time, as well as to multiple disjoint molecular fragments that share electrons. Thanks to these properties, ASTRA promises to appreciably expand the capabilities made available by existing molecular ionization codes.

This paper is organized as follows. Section II gives a short overview of the theoretical methods currently available or under active development to describe the ionization of polyelectronic molecular systems, and where ASTRA fits in this context. Section III defines the close-coupling (CC) states [26], the expressions for arbitrary operators within the TDMCC formalism, and relevant observables. Section IV illustrates how TDMCC is numerically implemented in ASTRA. Section V presents ASTRA results for the boron atom, the nitrogen molecule, and formaldehyde. Section VI offers conclusions and perspectives. Appendix A summarizes second-quantization formalism. Appendix B introduces transition-density matrices and their spin-adapted reduced counterparts. Appendix C details the derivation of the matrix elements between spin-adapted close-coupling states. Finally, Appendix D summarizes the CI-singles approximation.

---

*luca.argenti@ucf.edu

II. CONTEXT

A broad group of many-body methods for atomic and molecular ionization assumes a single reference for both the initial bound state and the final state in the

continuum. These include, the XATOM [27] and XMOLECULE codes [28], suited for estimating fragment yields following multiple sequential ionization steps by intense x-ray free-electron lasers (XFEL) pulses [29]; the E-POLYSCAT code [30,31], which has been applied to single-photon and multiphoton single ionization in the static-exchange approximation; TIRESIA B-spline density-function method, based on single center [32,33] and polycentric expansion [34,35], which was applied to strong-field ionization [36], core photoionization [37–39], pump-probe ionization of molecules with biological relevance [21,40], and high-harmonic generation [41]; OCTOPUS, a general-purpose time-dependent density-function theory (TDDFT) grid method applicable also to attosecond photoelectron spectroscopy [42] and transient-absorption spectroscopy [43]. While single-reference methods reproduce the essential features of strong-field ionization processes [44] of core photoelectron spectroscopy [45] and of processes that lead to the complete fragmentation of a target [46], wave-function-based multireference approaches are needed for those cases in which multiply excited states and the entanglement between photofragments play a fundamental role.

Time-dependent (TD) multiconfiguration (MC) self-consistent-field (SCF) approaches offer a compact description for the wave function. These include multiconfiguration TD Hartree [47] and Hartree-Fock [48–50], TD restricted-active-space (RAS) SCF [51,52], TD generalized-active-space (GAS) SCF [53,54], TD complete-active-space (CAS) SCF [55], TDMCSCF [56], and TD coupled-cluster methods [57]. These methods are effective for optical observables, whereas photoelectron observables are more difficult to converge since parent ions evolve even in the asymptotic limit, in absence of fields.

The remaining methods rely on the close-coupling approach (CC), which guarantees the correct asymptotic behavior of the wave function, and they give rise to a stable linear evolution. These methods differ in the level of accuracy with which the ions are represented, in the approximations made to evaluate the matrix elements between CC states, and in the limits imposed to the size and energy of the system.

In its simplest configuration-interaction singles (CIS) formulation [58–60], the CC space is described by single excitations from a reference single-determinantal function.

More general close-coupling approaches include correlation in the $(N-1)$-electron parent ions, as well as localized N-electron configurations to reproduce the short-range correlation of the neutral system in either bound or continuum states. Furthermore, different methods adopt different ways of reconciling the convenience of Gaussian-type orbitals (GTO) to describe bound states with the need of employing numerical bases to represent diffuse Rydberg and continuum orbitals, such as monocentric [61] and polycentric [34] B splines, or monocentric [62] and polycentric FEDVR functions [63].

The R-matrix method for molecular ionization [64], originally implemented for electron scattering and stationary regimes (UKRmol) [65], has recently been extended to time-dependent photoionization processes (UKRmol+) [66–68]. UKRmol+ employs hybrid Gaussian B-spline bases, and it can deal with arbitrary hybrid integrals. The K-matrix RPA approach has been applied to one- and two-photon transitions in stationary regime [69,70].

The complex-Kohn approach, based on the MESA program, (CK-MESA) has the unique feature of representing the effect of short-range correlation by means of an optical potential from a large CI space, which leads to accurate electron scattering and one-photon photoionization amplitudes [71–75]. Recently, CK-MESA has been extended to two-photon transitions and applied to resonant pump-probe process [76]. The time-dependent recursive indexing code (T-RECX) is a general software for the *ab initio* solution of large-scale time-dependent problems in atomic and molecular physics, including close coupling in a hybrid basis [77,78].

The XCHEM code circumvents the need of computing some exchange terms in the close-coupling Hamiltonian and some hybrid bielectronic integrals by confining the molecular ion in a spherical region in which all electronic states are expanded in terms of GTOs [79–85]. The polycentric B-spline basis developed for the TIRESIA code has recently been coopted for the development of a novel molecular ionization code based on the algebraic diagrammatic construction (ADC) to represent correlated many-body states [86,87]. The multi-channel Schwinger configuration interaction method (MCSCI), an established method for molecular photoionization, based on the graphical unitary group approach [88,89], has also been applied to the description of time-resolved photoelectron spectroscopy in linear molecules [90].

ASTRA belongs to the latter group of most general CC methods. It differs from the methods above because it relies on a TDMCC formalism. In ASTRA, the intricacies of the many-body interactions within the parent ions and between the ions and the photoelectron(s) are confined to the n-point correlation functions of highly correlated ionic states, which are efficiently computed by a large-scale-CI bound-state code. The computational cost for CC observables is independent of the size of the CI space for the ion. Exchange terms are treated exactly, so that the size of the target and the photoelectron energy are not constrained. ASTRA predictions are in excellent agreement with theoretical and experimental benchmarks for atoms and small molecules.

III. TDMCC FORMALISM

A. The close-coupling ansatz

In single photoionization processes, an initially bound molecular system M breaks into a photoelectron and a bound state A of the M$^+$ molecular parent ion. In the case of the absorption of a single photon γ from the ground state, for example, the process reads

$$M_g + \gamma \longrightarrow M_A^+ + e^-. \tag{1}$$
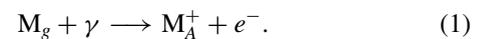
Once the two charged fragments are far apart, the parent ion can be found in only a finite number of states, i.e., all those permitted by conservation of energy, $E_A - E_g < \hbar\omega_\gamma$, where ω_γ is the photon frequency. When the parent ion and electron are well separated, the configuration space may be represented by a linear combination of parent ions times arbitrary one-particle states for the photoelectron,

$$\Psi \simeq \hat{\mathcal{A}}_N \sum_{A,P} c_{A,P} [\Phi_A(\mathbf{x}_1 \cdots \mathbf{x}_{N-1}) \varphi_P(\mathbf{x}_N)], \tag{2}$$

where $\mathcal{A}_N$ is the antisymmetrization operator [91], $\varphi_P(\mathbf{x}_N)$ is a spin orbital for the Nth electron, $\Phi_A(\mathbf{x}_1 \cdots \mathbf{x}_{N-1})$ is an ionic

state, $\mathbf{x} = (\vec{r}, \zeta)$ indicates an electron's spatial and spin coordinates, and $c_{A,P} \in \mathbb{C}$ are expansion coefficients. The *ansatz* (2), known as close-coupling (CC) approximation [26], accurately represents bound and autoionizing Rydberg series and multichannel scattering states for large values of the photoelectron radial coordinate.

At short range, a truncated CC expansion is able only in part to reproduce the correlated dynamics of the parent ion and the photoelectron. The CC approximation can be improved by including N-electron short-range configurations, which can be generated by extending the CC expansion to closed channels with thresholds well above the energy of interest. If the additional closed channels correspond to realistic ionic states, this procedure fails to account for the contribution of the ionic continuum, which may be relevant. A more effective and rapid convergence can be obtained by artificially confining the ionic states to a finite set of localized orbitals. In this paper, we follow this latter approach.

B. CC-matrix elements

This section defines single-ionization CC states and derives the matrix elements of arbitrary spin-free one-body and two-body operators between them. The CC partial-wave channel state formed by an ionic state A and a normalized spin orbital P, to represent the photoelectron, is

$$\Psi_{A,P}(\mathbf{x}_1, \ldots, \mathbf{x}_N)$$
$$= (-1)^{N-1}\sqrt{N}\hat{\mathcal{A}}_N \Phi_A(\mathbf{x}_1, \ldots, \mathbf{x}_{N-1})\,\phi_P(\mathbf{x}_N). \quad (3)$$

In this paper, it is convenient to employ the formalism of second quantization, in which the partial-wave states are expressed as

$$\Psi_{A,P}(\mathbf{x}_1, \ldots, \mathbf{x}_N) \;\longleftrightarrow\; |A, P\rangle = a_P^\dagger|A\rangle. \quad (4)$$

For a summary of the second-quantization formalism used in the present paper, see Appendix A.

To determine the bound and scattering electronic states of a molecule, and the electron dynamics induced by the exchange of photons with an external field, it is necessary to evaluate the matrix elements of the Hamiltonian and dipole operators between CC states. In second-quantization formalism, operators are written as combinations of strings of creators and annihilators. One-body operators are (see ch. 1 in [92])

$$\hat{O} = o_{RS}\, a_R^\dagger a_S, \quad (5)$$

where the sum over repeated indexes is implied, and $o_{RS} = \langle R|\hat{o}|S\rangle$ is the operator matrix element between spin orbitals for a single particle. The electron-electron repulsive Coulomb potential, a two-body operators, is

$$\hat{G} = \tfrac{1}{2}[PQ|RS]\, a_P^\dagger a_R^\dagger a_S a_Q \quad (6)$$

where $[PQ|RS]$ is the electrostatic repulsion term between the charge distributions $\rho_{PQ}(\vec{r}) = \sum_\zeta \phi_P^*(\vec{r}, \zeta)\phi_Q(\vec{r}, \zeta)$ and $\rho_{RS}(\vec{r}) = \sum_\zeta \phi_R^*(\vec{r}, \zeta)\phi_S(\vec{r}, \zeta)$,

$$[PQ|RS] = \int d^3r_1\, d^3r_2\, \rho_{PQ}(\vec{r}_1)\, r_{12}^{-1}\, \rho_{RS}(\vec{r}_2).$$

The calculation of the Hamiltonian and dipole matrix elements between CC states is thus reduced to evaluating strings

of up to six creation and annihilation operators between ionic states, e.g.,

$$\langle A, P|\hat{G}|B, Q\rangle = \tfrac{1}{2}\,[TV|RS]\,\langle A|a_P a_T^\dagger a_R^\dagger a_S a_V a_Q^\dagger|B\rangle.$$

Using the permutation symmetry of creation and annihilation operators, these operator strings can be written as a linear combination of strings in normal order (creators to the left, annihilators to the right). The matrix elements between ionic states of operators strings in normal order are known as transition-density matrices (TDM). The one-, two- and three-body TDMs are identified by the following notation [93],

$$\rho_{Q,P}^{BA} \equiv \langle A|a_P^\dagger a_Q|B\rangle,$$
$$\pi_{RS,PQ}^{BA} \equiv \langle A|a_P^\dagger a_Q^\dagger a_S a_R|B\rangle, \quad (7)$$
$$\gamma_{STU,PQR}^{BA} \equiv \langle A|a_P^\dagger a_Q^\dagger a_R^\dagger a_U a_T a_S|B\rangle.$$

As shown in Appendix A, using these definitions it is possible to compute matrix elements between single-ionization CC states for different operators by combining the TDMs with the operator matrix elements between basis orbitals. The overlap between CC states, for example, reads

$$\langle A, P|B, Q\rangle = \langle P|Q\rangle\delta_{AB} - \rho_{PQ}^{BA}. \quad (8)$$

Even if the overlap between different ionic states is diagonal by construction (the ionic states are eigenstates of the same model Hamiltonian), $\langle A|B\rangle = \delta_{AB}$, the overlap between the orbitals used to augment the ionic states to form the CC states, in general, is not: $\langle P|Q\rangle = s_{pq}\delta_{\pi\theta}$, where $\langle \mathbf{x}|P\rangle = \varphi_p(\vec{r})\,{}^2\chi_\pi(\zeta)$, $\langle \mathbf{x}|Q\rangle = \varphi_q(\vec{r})\,{}^2\chi_\theta(\zeta)$, $s_{pq} = \int d^3r\, \varphi_p^*(\vec{r})\varphi_q(\vec{r})$. The spin orbital $\langle \mathbf{x}|P\rangle$ is defined as the product of a spatial square-integrable function $\varphi_p(\vec{r}) \in \mathcal{L}^2(\mathbb{R}^3)$ and the spin wave function ${}^2\chi_\pi(\zeta)$, where 2 is the electron multiplicity and $\pi \in \{-\tfrac{1}{2}, \tfrac{1}{2}\}$ is the spin projection. Thanks to this generality, the TDMCC formalism is compatible with nonorthogonal numerical orbitals such as B splines [61] and FEDVR functions [94], which are ideally suited to represent Rydberg and photoelectron wavepackets. The matrix elements of one-body operators contain both one- and two-body TDMs,

$$\langle A, P|\hat{O}|B, Q\rangle = \delta_{AB}\, o_{PQ} + \langle P|Q\rangle\, O_{AB} - \rho_{P,R}^{BA}\, o_{RQ}$$
$$- o_{PS}\, \rho_{S,Q}^{BA} + o_{RS}\, \pi_{SP,QR}^{BA} \quad (9)$$

where $O_{AB} = o_{RS}\, \rho_{S,R}^{BA}$ is a matrix element between ionic states. Finally, two-body operators require TDMs up to the third order,

$$\langle A, P|\hat{G}|B, Q\rangle = \langle P|Q\rangle\, G_{AB} + [PQ|RS]\rho_{S,R}^{BA} - [PS|RQ]\rho_{S,R}^{BA}$$
$$+ [PT|RS]\pi_{TS,RQ}^{BA} + [QT|RS]\pi_{PS,RT}^{BA}$$
$$- \tfrac{1}{2}[TU|RS]\gamma_{USP,TRQ}^{BA} \quad (10)$$

where $G_{AB} = \tfrac{1}{2}[TU|RS]\pi_{US,TR}^{BA}$.

Expressions (8)–(10) are valid for any value of the spin quantum numbers of the excited orbitals and the ion states. Since in the present paper we restrict our attention to spin-free operators; however, it is convenient to determine the matrix elements between spin-adapted CC states (SACC), $|A, p; S\Sigma\rangle$, in which the spin of the ion and of the additional electron are

coupled to a well-defined spin S and spin projection Σ,

$$|A, p; S\Sigma\rangle = \sum_{\Sigma_A \pi} C^{S\Sigma}_{S_A \Sigma_A, \frac{1}{2}\pi} |A_{\Sigma_A}, p_\pi\rangle$$

$$= \sum_{\Sigma_A \pi} C^{S\Sigma}_{S_A \Sigma_A, \frac{1}{2}\pi} a^\dagger_{p_\pi} |A_{\Sigma_A}\rangle \quad (11)$$

where we have expanded the capital-letter index P into the (lower case) orbital index p and spin projection (Greek letter) π, $C^{S\Sigma}_{S_A, \Sigma_A, \frac{1}{2}\pi}$ are Clebsch-Gordan coefficients [95], with S_A and Σ_A being the ion spin and spin projection quantum numbers, respectively. The matrix element of a spin-free operator between SACC states, of course, vanishes unless the two states have the same spin S and spin projection Σ. For brevity, therefore, in the following we will adopt the shorthand notation

$$\langle A, p|\hat{M}|B, q\rangle \equiv \langle A, p; S\Sigma|\hat{M}|B, q; S\Sigma\rangle. \quad (12)$$

The matrix elements between SACC states are derived in Appendix A and summarized below. As explained more in detail in the next section, it is convenient to partition the spatial orbitals in inactive or core (I), active (A), virtual (V), and external (E) orbitals. The TDMs vanish unless all their indexes correspond to inactive or active orbitals. Furthermore, the expressions for the TDMs with inactive indexes can be simplified in such a way that the TDMs need to be evaluated only for active-orbital indexes. In the following, all sums over orbital indexes appearing in TDMs are assumed to be limited to active orbitals only (or to vanish otherwise). The overlap, for example, reads

$$\langle A, p|B, q\rangle = \delta_{AB} s_{pq} + W^{BA}_{pq}, \quad (13)$$

where W^{BA}_{pq} is a reduced one-body TDM obtained by coupling the spin projections of $\rho^{BA}_{P,Q}$ (see Appendix C). In (13), the W^{BA}_{pq} vanishes unless both p and q are active, in which case $s_{pq} = \delta_{pq}$. In all the other cases, only the $\delta_{AB} s_{pq}$ term survives. For a generic one-body operator $\hat{O}$,

$$\langle A, p|\hat{O}|B, q\rangle = \delta_{AB} o_{pq} + s_{pq} O_{AB} + \langle \hat{O}\rangle_{\text{core}} W^{BA}_{pq}$$
$$+ W^{BA}_{pr} o_{rq} + o_{ps} W^{BA}_{sq} + o_{rs} P^{BA}_{rs,qp}. \quad (14)$$

See Appendix C for the definition of the reduced TDM $P^{BA}_{rs,qp}$ in terms of $\pi^{BA}_{PS,RQ}$. The term $\langle \hat{O}\rangle_{\text{core}} = 2o_{xx}$ represents the contribution to the observable from the ionic inactive orbitals. Finally, the complete Hamiltonian matrix element between SACC states comprises both mono- and bielectronic components and it involves up to three-body TDMs. The Hamiltonian matrix elements between CC states that do not contain any virtual or external orbital are most readily evaluated using the same quantum-chemistry program that computes the TDMs. For this reason, the current implementation of ASTRA does not explicitly use three-body TDMs. The general matrix element between SACC in which at least one of the two states involves a virtual or external orbital, reads

$$H_{Ap,Bq} = \langle A, p|\hat{H}|B, q\rangle$$
$$= \delta_{AB}(s_{pq}E_A + \tilde{h}_{pq}) + W^{BA}_{pr}\tilde{h}_{rq} + \tilde{h}_{pr}W^{BA}_{rq}$$
$$+ \delta_{S_A S_B}[pq|rs]Q^{BA}_{sr} + [ps|rq]W^{BA}_{sr}$$
$$+ [pt|rs]P^{BA}_{ts,qr} + [qt|rs]P^{BA}_{ps,tr} \quad (15)$$

where E_A is the energy of the parent ion A, and where the monoelectronic matrix element h_{pq} has been replaced by an effective Hamiltonian $\tilde{h}_{pq}$ that incorporates the Coulomb and exchange terms with the core electrons,

$$\tilde{h}_{pq} = h_{pq} + (2[pq|\tilde{r}\tilde{r}] - [p\tilde{r}|\tilde{r}q]). \quad (16)$$

The reduced TDM Q^{BA}_{sr} is defined in Appendix C [see (C8)].

C. Observables

In this paper, we focus on the parameters of bound states and autoionizing states of selected atomic and molecular systems, as well as on the total photoionization cross section of each system from its ground state. The energy of the bound states and the position and width of the autoionizing states are obtained by diagonalizing the electronic fixed-nuclei Hamiltonian in the SACC basis described above, with outgoing boundary conditions. The radial functions used to build the CC channels are confined to a quantization box $r < R_{\text{box}}$, with typical size of the order of few hundred Bohr radii. Outgoing boundary conditions are enforced by adding to the Hamiltonian a complex absorption potential (CAP) [96] V_{CAP} with support located near the boundary of the quantization box,

$$\tilde{H} = \sum_i \left[\frac{p_i^2}{2} - \sum_A \frac{Z_A}{|\vec{r}_i - \vec{R}_A|} \right] + \sum_{i>j} \frac{1}{|\vec{r}_i - \vec{r}_j|}$$
$$+ \sum_{A>B} \frac{Z_A Z_B}{|\vec{R}_A - \vec{R}_B|} + V_{\text{CAP}}, \quad (17)$$

$$V_{\text{CAP}} = -i \sum_{\alpha i} b_\alpha \theta(r_i - R_\alpha)(r_i - R_\alpha)^2, \quad (18)$$

where Z_A and $\vec{R}_A$ are the charge and position of nucleus A, while $\vec{r}_i$ and $\vec{p}_i$ are the position and momentum operator of the ith electron. Here, the CAP is parametrized as the sum of several negative imaginary parabolic potentials, $\alpha = 1, 2, \ldots, N_{\text{CAP}}$, each starting at a prescribed radius R_α; $\theta(x)$ is the Heaviside step function, and b_α are real positive constants. In dynamical conditions, the CAP prevents artificial reflections of a photoelectron wavepacket from the box boundary. While local multiplicative CAP are not perfect absorbers [96], the reflection coefficient of an electron from either the box boundary or the CAP itself can easily be reduced to machine precision in any energy range of interest by combining multiple parabolic potentials across a sufficiently large radial distance. For the present paper, in which we examine total energies within one Hartree from the ionization threshold, we found that a combination of a broad and shallow complex potential ($b_1 \simeq 10^{-6}$, $R_1 \simeq R_{\text{box}} - 200$ a.u.), which efficiently absorbs slow electrons without reflecting them, and a narrower and steeper complex potential ($b_2 \simeq 10^{-4}$, $R_2 \simeq R_{\text{box}} - 100$ a.u.), which absorbs fast electrons before they get reflected by the box boundary, is sufficient to achieve convergence.

Let $|\boldsymbol{\phi}\rangle = (|\phi_1\rangle, |\phi_2\rangle, \ldots)$ be the CC basis and $\tilde{\mathbf{H}} = \langle \boldsymbol{\phi}|\tilde{H}|\boldsymbol{\phi}\rangle$ be the representation of the Hamiltonian with CAPs in this basis. The diagonalization of $\tilde{\mathbf{H}}$ can be written as

$$\tilde{\mathbf{H}} = \mathbf{U}_R \tilde{\mathbf{E}} \mathbf{U}_L^\dagger, \quad (19)$$

where $\tilde{\mathbf{E}}_{ij} = \delta_{ij}\tilde{E}_i$ is the diagonal matrix of the complex eigenvalues of the projected $\tilde{H}$, whereas $\mathbf{U}_{L/R}$ are the left/right eigenvectors, normalized so that $\mathbf{U}_L^\dagger \mathbf{U}_R = \mathbf{1}$. The complex energies thus obtained can be grouped in three types: (i) those below the first ionization threshold, with negligible imaginary part, which correspond to bound states that have negligible amplitude in the CAP region; (ii) sequences of eigenvalues that branch out from each threshold and rapidly acquire large negative imaginary components, which correspond to nonresonant continuum states; and (iii) isolated complex eigenvalues $\tilde{E}_i = \bar{E}_i - i\Gamma_i/2$, where $\bar{E}_i = \Re e(\tilde{E}_i)$ and $\Gamma_i = -2\Im m(\tilde{E}_i)$, which are largely independent on the choice of the extinction parameter b_α and represent the complex energies of autoionizing states. In the region where the CAP vanishes, $\forall i,\ \alpha,\ r_i \leqslant R_\alpha$, the right eigenstates of $\tilde{\mathbf{H}}$ can be regarded as Siegert states [97–99].

The spectral resolution of the Hamiltonian in terms of states that satisfy outgoing boundary conditions also allows us to compute the total photoionization cross section, which in length gauge reads

$$\sigma_{\text{tot}}(\omega) = \frac{4\pi^2\omega}{c} \sum_\alpha \left| \langle \Psi_{\alpha E_g + \omega} | \hat{\epsilon} \cdot \vec{\mu} | g \rangle \right|^2, \tag{20}$$

where $|g\rangle$ is the ground state of the target, $c \simeq 137.035$ is the speed of light in atomic units, $\hat{\epsilon}$ is the light-polarization unit vector, $\hat{\epsilon}^* \cdot \hat{\epsilon} = 1$, $\vec{\mu} = -\sum_{i=1}^{N_e} \vec{r}_i$ is the electronic dipole moment, and $|\Psi_{\alpha E}\rangle$ are a complete set of orthonormal single-ionization scattering states for all channels α open at the energy $E_g + \omega$, $\langle \Psi_{\alpha E} | \Psi_{\beta E'} \rangle = \delta_{\alpha\beta}\delta(E - E')$. Indeed, for any localized wavepacket $|\phi\rangle$,

$$\sum_\alpha |\langle \Psi_{\alpha E}|\phi\rangle|^2 = \sum_\alpha \langle \phi | \Psi_{\alpha E} \rangle \langle \Psi_{\alpha E} | \phi \rangle$$
$$= \langle \phi | \delta(E - H) | \phi \rangle$$
$$= -\frac{1}{\pi} \Im m[\langle \phi | G^+(E) | \phi \rangle], \tag{21}$$

where $G^+(E) = (E - H + i0^+)^{-1}$ is the retarded resolvent [100]. The total photoionization cross section, therefore, can be written as ($|\phi\rangle = \hat{\epsilon} \cdot \vec{\mu}|g\rangle$)

$$\sigma_{\text{tot}}(\omega) = -\frac{4\pi\omega}{c}\Im m\langle g| \hat{\epsilon} \cdot \vec{\mu}\ G_0^+(E_g + \omega)\ \hat{\epsilon} \cdot \vec{\mu}\ |g\rangle, \tag{22}$$

which is nothing other than the optical theorem applied to photoionization processes. Since wavepackets of the form $G^+(E)|\phi\rangle$ contain outgoing components only [100,101], we can use the numerical spectral resolution of $\tilde{H}$ to write down an explicit expression for the retarded resolvent in the CC basis,

$$\mathbf{G}_0^+(E) = \mathbf{U}_R \frac{1}{E - \tilde{\mathbf{E}}} \mathbf{U}_L^\dagger. \tag{23}$$

Equation (23) is accurate as long as the resolvent is applied to wavepackets localized in the region where the CAP vanishes and the result is evaluated in the same region. This is the case for the dipole operator acting on the ground state of the system, and hence

$$\sigma_{\text{tot}}(\omega) = -\frac{4\pi\omega}{c}\Im m[\boldsymbol{\mu}^\dagger \mathbf{G}_0^+(E_g + \omega)\boldsymbol{\mu}], \tag{24}$$

where $\boldsymbol{\mu} = \langle \boldsymbol{\phi} | \hat{\epsilon} \cdot \vec{\mu} | g \rangle$. This expression bypasses the calculation of scattering states and it can be evaluated at a negligible computational cost. The calculation of multichannel scattering states is beyond the scope of the present paper and will be examined in future works.

IV. ASTRA IMPLEMENTATION

In this section, we examine some aspects of the numerical implementation of ASTRA.

To reproduce molecular ionization observables, ASTRA relies on internal components as well as on external tools: DALTON [102], LUCIA [103], and GBTOlib [67]. DALTON is a general-purpose quantum-chemistry program used to compute an initial set of GTOs as well as mono- and bi-electronic integrals between them. LUCIA is a molecular electronic-structure program with a focus on efficient large-scale configuration interactions [103]. Starting from an initial set of orbitals and integrals, LUCIA can carry out sophisticated calculations to optimize the initial orbital, compute correlated bound states, as well as the TDMs between them (see Sec. IV A for details). GBTOlib is a public library, part of the UKRmol+ molecular scattering package that can compute electronic integrals between hybrid Gaussian/B-spline functions. ASTRA employs also an additional independent set of B-spline spherical functions external to the molecular region and whose support is effectively disjoint from that of any GTO orbitals. The flow diagram in Fig. 1 illustrates how the different external and internal components of ASTRA are connected with each other. The external components are managed by a setup script. One interface (A TDM IF) converts the TDMs generated by LUCIA to their reduced spin-adapted form. A second interface (A INT IF) converts the hybrid integral generated by GBTOlib to the internal database of ASTRA and complements the hybrid integrals with those involving the additional set of B splines that extends beyond the molecular region.

The right frame of the flow diagram delimits the core section of ASTRA. The ASTRA operator program (A OP) generates the representation of arbitrary operators in the SACC space, starting from the reduced TDM, the hybrid integrals, and the specifics of the close-coupling configuration file (not shown). Subsequently, the ASTRA TISE program determines the eigenstates and eigenvalues of the system with either vanishing or outgoing boundary conditions. In parallel, the Lippmann-Schwinger equation solver (A LSE) determines the scattering states. The generalized and the confined eigenstates of the system can then be used to evaluate stationary observables, such as the one-photon total ionization cross section (see Sec. V), the photoelectron emission due to external pulses (A TDSE, A PSS), by solving the TDSE in the full CC basis, or the optical response by solving the TDSE in a reduced model space of Siegert states (A ES TDSE, A OPT RES). Due to its reliance on quantum-chemistry external tools, ASTRA only handles molecular spatial symmetry groups that are subgroups of D_{2h} [91].

A. CASCI ionic states and TDM

The molecular orbitals can be determined using LUCIA for several wave-function methods: Hartree-Fock, complete
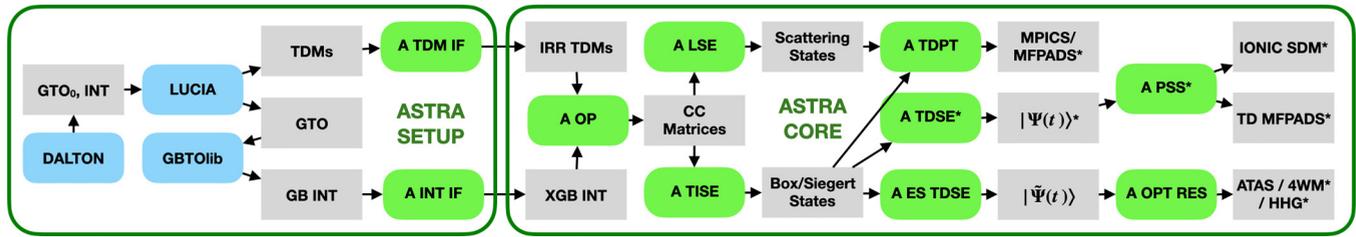
FIG. 1. ASTRA flow scheme. Rounded boxes represent executables, whereas square boxes represent databases. The "A" prefix indicates an executable from the ASTRA suite. In the setup section (left frame), from a set of configuration files (not shown), the ASTRA setup script invokes, in sequence: (i) DALTON, to compute a starting set of GTO and their associated integrals; (ii) LUCIA, to optimize the orbitals, to compute the correlated ionic states, as well as the TDM between them; (iii) GBTOlib, to compute, from the optimized orbitals, the hybrid integrals with the B splines in the molecular region. See text for details on the steps that involve only ASTRA components. Starred elements are currently under development.

active space self-consistent field calculations (CASSCF), and generalized and restricted-active-space SCF calculations (GASSCF and RASSCF). Molecular ionization processes start from a neutral state and lead to the production of multiple ionic states with different symmetries and multiplicities. It is therefore advantageous to optimize the orbitals so they provide a balanced representation of a set of target states with different number of electrons, multiplicities, spatial symmetry, and energy. Such optimizations go under the name of state-averaged (SA) MCSCF calculations, and LUCIA is capable of optimizing the orbitals for an ensemble of such states. This makes LUCIA a well-suited QC engine for generating the orbitals and CI expansions that subsequently are used in the photoionization codes.

In a subsequent step, LUCIA computes the required TDMs between all the ionic states in the CC expansion, as well as the overlap, Hamiltonian, and dipole matrix elements between all the states obtained by augmenting these ions by any of the active orbitals. For one-electron photoionization processes, one- and two-electron density matrices are sufficient, but LU-CIA can also calculate the three-electron TDMs required for double photoionization. The TDMs are obtained in LUCIA over spin orbitals and are subsequently spin-adapted by the ASTRA interface. The LUCIA code determines the TDMs from CI expansions in the Slater determinant (SD) basis. However, in order to assure that the involved states have well-defined spin of the involved states, these states are defined in terms of spin-adapted configuration state functions (CSFs). The calculations of the TDMs and the inner part of a direct CI step are thus preceded by a transformation from the CSF to the SD basis.

The expansion of a CI state in terms of Slater determinants is realized using the formalism of spin strings [92,103,104]. In this formalism, a SD is expressed as an alpha-spin string, which is an ordered product of the creation operators for the alpha-spin orbitals, times a beta-spin string, which is an ordered product of creation operators for beta-spin orbitals. The use of spin strings provides a formally simple approach that can perform the CI optimization and the subsequent calculation of TDMs in an efficient manner. An advantage of the spin-string formalism is that the information required to determine the action of a general operator on a CI expansion is easy to determine and store. In the present implementation, only information about single-electron removal and addition

from strings is determined and stored. The action of operators containing several creation or annihilation operators is constructed on the fly from the lists of one-electron removals and additions.

The TDMs are also written and computed in terms of spin strings [105,106]. A TDM is thus written as matrix with indices defining spin strings,

$$\rho_{\nu,\mu}^{BA} = \langle A | \mathcal{O}_\mu^\dagger \mathcal{O}_\nu | B \rangle. \quad (25)$$

By letting $\mathcal{O}$ be strings of one-, two-, and three-electron annihilation operators, the one-, two- and three-electron density matrices of Eq. (7) are obtained. The various forms of the TDMs are calculated using a common suite of routines, where the states $\mathcal{O}_\mu | A \rangle$, $\mathcal{O}_\nu | B \rangle$ first are determined as expansions in SDs $|I\rangle$,

$$\mathcal{O}_\mu | A \rangle = \sum_I C_{\mu I}^A |I\rangle, \quad \mathcal{O}_\nu | B \rangle = \sum_I C_{\nu I}^B |I\rangle. \quad (26)$$

For a given state A, the dimension of $C_{\mu I}^A$ and the operation count for its evaluation is approximately $N^n N_{det}$, where N is the number of active orbitals, n is the order of the TDM, and N_{det} is the number of SDs in the expansion. In a second step, the TDM is obtained by matrix multiplication,

$$\rho_{\nu,\mu}^{BA} = \langle A | \mathcal{O}_\mu^\dagger \mathcal{O}_\nu | B \rangle = \sum_I C_{\mu I}^{A*} C_{\nu I}^B. \quad (27)$$

This matrix multiplication scales as $N^{2n} N_{det}$. For an example with 20 active orbitals and one million SDs, the operation count for the evaluation of $C_{\nu I}^B$ for a two-electron TDM is about 4×10^8 and the following matrix multiplication requires about 3×10^{11} operations. In the above estimates, simplifications arising from spatial symmetry were neglected. For a molecule with N_{IRR} irreducible representations, the operation count the the matrix multiplication is reduced by a factor of about N_{IRR}^2. On a single core of a modern CPU, the discussed TDM for a molecule with four irreducible representations may thus be evaluated in a few seconds.

B. Orbital basis

As commented in Sec. III B, the expressions for the CC-matrix elements suggest a natural partition of the orbitals in inactive, active, virtual, and external, which helps simplify
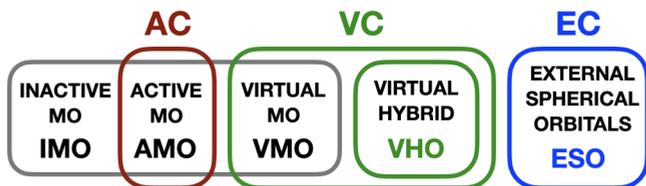
FIG. 2. Schematic classification of the electronic orbitals used in ASTRA CC expansion. The labels over the frames, AC, VC, and EC, indicate the active, virtual, and external component of the close-coupling channel obtained by augmenting a given ion by the full set of the noninactive orbitals.

their calculation. Figure 2 illustrates schematically the different orbitals.

An initial set of atomic GTO orbitals are used to generate an orthonormal set of molecular orbitals (MO), ϕ_i^M, $\langle \phi_i^M | \phi_j^M \rangle = \delta_{ij}$, optimized by means of a self-consistent-field calculation (HF, RASSCF, CASSCF, etc.). All MOs are negligible beyond a certain distance R_{mol} from the molecular barycenter. We call $r < R_{mol}$ the molecular region. The MOs are divided in inactive orbitals (IMO), which are doubly occupied in all the ions, active orbitals (AMO), which have variable occupation numbers in the ions, and virtual orbitals (VMO), unoccupied in the ions,

$$\{\phi_i^M\} = \{\phi_i^{IM}\}_{i \in I_{IM}} \cup \{\phi_i^{AM}\}_{i \in I_{AM}} \cup \{\phi_i^{VM}\}_{i \in I_{VM}}$$

The MOs are assumed to be adequate to represent the correlated state of the $N-1$ electrons of the ion. However, they are normally insufficient to describe the state of an Nth electron in interaction with the ion, particularly for energies close to or above the ionization continuum. The virtual hybrid orbitals (VHO) are obtained by complementing the virtual MOs with a set of internal spherical B splines [61,107,108], with support within the molecular region,

$$\chi_{n\ell m}^{int}(\vec{r}) = N_n r^{-1} B_n^{int}(r) X_{\ell m}(\hat{r}),$$ (28)

where $X_{\ell m}(\hat{r})$ are real spherical harmonics and N_n are normalization constants. The internal B splines are orthonormalized to the MOs,

$$\phi_i^{VH} = \sum_{nm\ell} \chi_{n\ell m}^{int} c_{nm\ell, i} + \sum_j \phi_j^M c_{ji},$$ (29)

$$\langle \phi_i^{VH} | \phi_j^M \rangle = 0, \quad \langle \phi_i^{VH} | \phi_j^{VH} \rangle = \delta_{ij}.$$ (30)

This second step is carried out by GBTOlib [67], by performing a Gram-Schmidt orthogonalization of the B-spline basis with the MOs (leaving the MOs unchanged) followed by a symmetric orthogonalization of the resulting set. The program also evaluates the monoelectronic and bielectronic integrals involving these hybrid orbitals.

The MOs and VHOs together afford to the CC space considerable flexibility for the description of a photoelectron subject to the polycentric field of the ion. Finally, the external spherical orbitals (ESO) are spherical B splines that extend the internal B spline set beyond the molecular region

$$\phi_{n\ell m}^E(\vec{r}) = N_n r^{-1} B_n^{ext}(r) X_{\ell m}(\hat{r}).$$ (31)
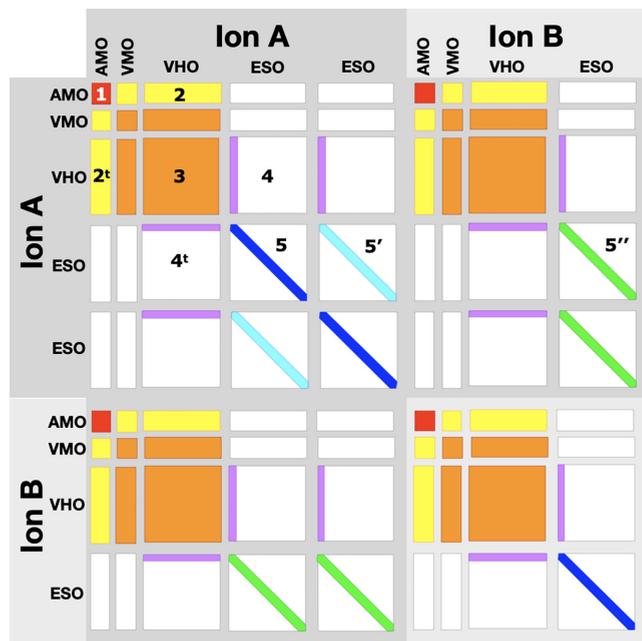
FIG. 3. Block structure of the Hamiltonian matrix between the CC states. The parts in white are empty. While the blocks are not to scale, it is clear that the matrix is highly sparse. There are essentially five different types of blocks (colored online), which correspond to as many separately optimized formulas and set of integrals: 1 for (C22), 2 for (C23), 3 for (C24) and for 4, 5 see (C25).

The MOs and the external B splines have effectively disjoint support

$$\phi_i^M(\vec{r}) \, \phi_{n\ell m}^E(\vec{r}) \simeq 0, \quad \forall \vec{r},$$ (32)

which means that the integral of any local operator between an external B spline and an hybrid orbital can be expressed in terms of integrals between B-spline functions only.

C. Matrix elements

Taking into account the different groups of orbitals used to form the CC channels, we can further simplify the expression for the matrix elements of the different operators given in Sec. III. The structure of the Hamiltonian matrix is schematized in Fig. 3, where the colored regions identify coupling between channels while the white areas correspond to regions with null matrix elements. The evaluation of the overlap and of one-body operators is relatively simple. The most complex monoelectronic matrix elements to evaluate are those between AC states, which involves both 1B-TDM and 2B-TDM [see (14)]. Overall, however, purely monoelectronic operators are straightforward to evaluate. In this section, therefore, we will focus on the Hamiltonian only. Apart for the complex absorbing potential, which affects exclusively the diagonal blocks between external orbitals, the rest of the Hamiltonian is Hermitian. Therefore, without loss of generality, we can limit the description of the Hamiltonian matrix elements to the upper triangular set of blocks of the close-coupling matrix. The evaluation of an Hamiltonian matrix element requires the 3B-TDM only when both CC states belong to an AC. Since this matrix elements involves MOs only, it is convenient to

carry out the calculation directly within LUCIA. ASTRA is then left with the minimal task of evaluating, from the Hamiltonian matrix element between CC states, only those involving at least a VC or EC state, as detailed in Appendix C 4.

Let us now consider the general expression (15) for the other blocks. For the active-virtual blocks (AV), the Hamiltonian is

$$H^{AV}_{Ap,Bq} = \delta_{AB}\tilde{h}_{pq} + W^{BA}_{pr}\tilde{h}_{rq} + [pq|rs]Q^{BA}_{sr}$$
$$+ [ps|rq]W^{BA}_{sr} + [qt|rs]P^{BA}_{ps,tr},$$

whereas the blocks between active and external states are zero, $H^{AE}_{Ap,Bq} = 0$. Notice that the AV blocks are the only ones involving 2B-TDMs. Since there are comparatively few such elements, the calculation is inexpensive. For the VV block, the Hamiltonian contains both local and nonlocal interchannel coupling potentials,

$$H^{VV}_{Ap,Bq} = \delta_{AB}(E_A\delta_{pq} + \tilde{h}_{pq}) + [pq|rs]Q^{BA}_{sr} + [ps|rq]W^{BA}_{sr},$$

whereas in the VE and EE blocks only the local potential remains (X = V, E),

$$H^{XE}_{Ap,Bq} = \delta_{AB}(E_A\delta_{pq} + \tilde{h}_{pq}) + [pq|rs]Q^{BA}_{sr}. \quad (33)$$

The local potential $[pq|rs]Q^{BA}_{sr}$ describes the repulsion between two disjoint distributions of charge, one confined within and one outside the molecular region, i.e., the multipolar electrostatic coupling between channels. The expression, therefore, can be conveniently rearranged by means of a multipole expansion of the electron-electron repulsion energy

$$\frac{1}{|\vec{r}-\vec{r}'|} = \sum_{\ell m}\frac{4\pi}{2\ell+1}\frac{r^{\ell}_<}{r^{\ell+1}_>}X_{\ell m}(\hat{r})X_{\ell m}(\hat{r}'), \quad (34)$$

where $X_{\ell m}(\hat{r})$ are real spherical harmonics [109]. The bielectronic matrix element $[pq|rs]$ then becomes

$$[pq|rs] = \sum_{\ell m}o^{(\ell m)}_{pq}M^{(\ell m)}_{rs},$$

where

$$o^{(\ell m)}_{pq} = \langle p|\frac{X_{\ell m}(\hat{r})}{r^{\ell+1}}|q\rangle, \quad M^{(\ell m)}_{rs} = \frac{4\pi\langle r|r^{\ell}X_{\ell m}(\hat{r})|s\rangle}{2\ell+1}.$$

We can now recognize that the one-body part of the Hamiltonian $\tilde{h}_{pq}$ defined in (16) already contains a multipolar nuclear attraction potential, whose associated set of moments are given by

$$M^{(\ell m)}_{\text{nuc}} = \frac{4\pi}{2\ell+1}\sum_{\alpha}^{N_{\text{at}}}Z_{\alpha}R^{\ell}_{\alpha}X_{\ell m}(\hat{R}_{\alpha}). \quad (35)$$

It is convenient, therefore, to gather the nuclear attraction and the electronic repulsion terms in a single-transition multipolar momentum $\mathcal{M}^{(\ell m)}_{AB}$, defined as

$$\mathcal{M}^{(\ell m)}_{AB} = \sum_{rs}'M^{(\ell m)}_{rs}Q^{BA}_{sr} + 2\sum_{x}^{\text{core}}M^{(\ell m)}_{xx} - M^{(\ell m)}_{\text{nuc}}, \quad (36)$$

The monopolar momentum, of course, is simply the ionic charge, $\mathcal{M}^{(00)}_{AB} = \delta_{AB}Z_{\text{ion}}$, and the expression for the Hamiltonian in the outer region given in (33) can therefore be rewritten as

$$H^{XE}_{Ap,Bq} = \delta_{AB}(s_{pq}E_A + t_{pq} - Z_{\text{ion}}o^{(00)}_{pq})$$
$$+ \sum_{\ell m}^{\ell>0}o^{(\ell m)}_{pq}\mathcal{M}^{(\ell m)}_{AB}. \quad (37)$$

where t_{pq} is a kinetic-energy matrix element.

V. RESULTS

This section illustrates ASTRA performance by comparing its results with well established and independent single-ionization codes, together with experimental references, when available. We apply ASTRA to the calculation of selected bound and resonant state parameters, and total photoionization cross section of three reference model systems: the boron atom, the nitrogen molecule, and formaldehyde, H_2CO, as a representative of an open-shell atomic system, a diatomic molecule, and a simple polyatomic molecule, respectively. The results are compared with the best values available in the literature, with which we find an excellent agreement. In the case of N_2, we consider different levels of approximation to illustrate both the consistency of the method with independent benchmark calculations, such as CIS, as well as the flexibility with which electronic correlation in ionization can be treated already within the current implementation of ASTRA.

A. Boron

In this section, we discuss the computation of the bound-state energies, resonance parameters, and photoionization cross-section of atomic boron and compare the results with the NEWSTOCK suite of atomic photoionization codes [110] as well as with a dedicated three-active-electron atomic code (TAEC) [111]. Boron doublet CC space is built from ions with both singlet and triplet multiplicity. This comparison, therefore, is particularly relevant because it allows us to test the correct implementation of the TDMCC method when TDMs between states with different multiplicity are required. Although both ASTRA and NEWSTOCK use equivalent CC expansions, the two approaches are intrinsically different. NEWSTOCK, as an atomic photoionization code, uses the full $SO(3)$ symmetry of the atom, whereas ASTRA uses the D_{2h} Abelian point group. NEWSTOCK relies on the nonrelativistic multiconfiguration Hartree-Fock (MCHF) program of the ATSP2K [112] package to compute the parent-ion states. To obtain the equivalent degree of local electron correlation, we employ the same number of ionic states in ASTRA and NEWSTOCK calculations. The ionic states are optimized by including all possible single and double excitations from the reference determinant with active orbitals up to the principal quantum number $n = 4$. TAEC [111] uses a virtually complete two-electron basis for the two active valence electrons of the B^+ ion in the field of a polarizable core, as well as an optimized set of several thousands configurations selected from the full-CI three-active-electron space, and a significantly larger CC expansions compared with either NEWSTOCK or AS-TRA. As one may expect from a dedicated program, therefore, TAEC results are closer to the experimental values listed in the NIST database [113]. In constructing the CC expansion for

TABLE I. Comparison of the energies of the first 11 singlet and triplet bound states of B II, $E - E_{2s^2}$ (eV).

Conf.	ASTRA	NEWSTOCK	TAEC [111]	Exp. [113]
$2s2p$ $(^3P^o)$	4.6563	4.6308	4.6391	4.6317
$2s2p$ $(^1P^o)$	9.3240	9.3421	9.1187	9.1000
$2p^2$ $(^3P^e)$	12.373	12.256	12.287	12.266
$2p^2$ $(^1D^e)$	12.856	12.860	12.706	12.691
$2p^2$ $(^1S^e)$	16.012	16.001	15.835	15.827
$2s3s$ $(^3S^e)$	16.036	16.037	16.095	16.089
$2s3s$ $(^1S^e)$	17.166	17.300	17.069	17.062
$2s3p$ $(^3P^o)$	17.839	17.831	17.858	17.853
$2s3p$ $(^1P^o)$	17.880	17.910	17.874	17.866
$2s3d$ $(^3D^e)$	18.652	18.650	18.681	18.678
$2s3d$ $(^1D^e)$	19.230	19.380	19.183	19.178

ASTRA and NEWSTOCK calculations, we have included the first 12 parent ions, which are listed in Table I. For the ASTRA computation, the ionic Hartree-Fock orbitals are generated with DALTON [102] using the aug-cc-pVQZ basis. In ASTRA and NEWSTOCK, the results of the ionic states with dominant configurations up to $2s3d(^1D^e)$ are compared in Table I relative to the $2s^2(^1S^e)$ ionic ground state. The agreement between our results and the experimental values is remarkable. The values computed with ASTRA lie within 0.03 to 0.2 eV of the experimental values, which is deemed sufficient to ascertain the consistency of ASTRA's implementation. In both ASTRA and NEWSTOCK, the twelve parent ions are coupled to photoelectrons with angular momentum up to $l_{max} = 3$, within a quantization box of 300 a.u., employing B splines of order seven and node separation of 0.3 a.u. In ASTRA, $R_{mol} = 20$ a.u. and a single CAP is used, which starts at $R_\alpha = 200$ a.u.

Table II lists the first few bound-state energies of B I with doublet and quartet symmetries. For the $2s^2 2p(^2P^o)$ ground state, the absolute value of energy obtained from ASTRA is $-24.600\,14$ a.u., which differs from more accurate nonrelativistic MCHF calculations [114] by 0.0538 a.u. This energy difference is commensurate with the value of 0.044 735 a.u estimated for the K-shell correlation energy [115], which is not accounted for in the present calculation. The residual discrepancy is compatible with the larger active space used in the MCHF calculation.

In ASTRA, the resonance parameters are obtained by diagonalizing the multichannel CC Hamiltonian in the presence of a CAP, whereas NEWSTOCK enforces outgoing boundary conditions using exterior-complex scaling (ECS) [116–118]. Table III compares the effective quantum number, $n^* = [2(E_{thr} - E_n)]^{-1/2}$, and resonance width Γ of selected

TABLE II. Bound-state energies of B I, $E - E_{2s^2 2p}$ (eV).

Conf.	ASTRA	NEWSTOCK	TAEC [111]	Exp. [113]
$2s2p^2$ $(^4P^e)$	3.584	3.564		3.552
$2s^23s$ $(^2S^e)$	4.869	4.881	4.958	4.964
$2s2p^2$ $(^2D^e)$	5.916	5.927	5.939	5.933
$2s^23p$ $(^2P^o)$	5.993	5.990	6.021	6.027
$2s^23d$ $(^2D^e)$	6.707	6.682	6.785	6.790

TABLE III. Comparison of effective quantum number, n^* and resonance width Γ (a.u.), of $^2P^o$ resonances between $2s^2$ and $2s2p(^3P^o)$ thresholds. The notation 5.94[–3] stands for 5.94×10^{-3}.

Conf.	ASTRA n^*	ASTRA Γ	NEWSTOCK n^*	NEWSTOCK Γ	TAEC [111] n^*	TAEC [111] Γ
$2s2p3s$	2.146	5.94[–3]	2.131	5.58[–3]	2.140	5.73[–3]
$2s2p4s$	3.120	1.79[–3]	3.157	1.63[–3]	3.157	1.59[–3]
$2s2p5s$	4.170	6.98[–4]	4.164	6.96[–4]	4.161	6.69[–4]
$2s2p6s$	5.172	3.59[–4]	5.166	3.60[–4]	5.162	3.44[–4]
$2s2p7s$	6.173	2.10[–4]	6.167	2.10[–4]	6.163	2.00[–4]

autoionizing states with $^2P^o$ symmetry, between the the $2s^2(^1S^e)$ and $2s2p(^3P^o)$ B$^+$ thresholds. The results for the dominant $2s2p(^3P^o)ns$ series from ASTRA are in excellent agreement with those obtained with NEWSTOCK and TAEC. Table IV shows the parameters for the first few terms of the $2s2p(^3P^o)np$ $(^2S^e)$ autoionizing states. The resonance parameters for the two main series, $2s2p(^3P^o)np$ and $2s2p(^3P^o)nf$ of $^2D^e$ autoionizing states are listed in Table V. Figure 4 compares the total photoionization cross-section of boron, from the $^2P^o$ ground state to the energy region between the $2s^2$ and the $2s2p$ B$^+$ thresholds, computed in the length gauge with ASTRA, with the one computed with NEWSTOCK, both in length and in velocity gauge. As discussed in Sec. III, in this paper, ASTRA evaluates the total cross section using the optical theorem (24), whereas the cross section in NEWSTOCK is computed from the dipole transition matrix elements between the ground state and a complete set of scattering states of the atom. The two calculations, which differ entirely in their methodology, are in remarkable agreement with each other. In particular, ASTRA accurately reproduces the characteristic asymmetry of the multichannel resonant profiles of the many autoionizing states in this region.

Figure 4(a) covers the whole energy interval between the first and second threshold. Figure 4(b) shows a close up on the resonant region, and identifies selected terms in the three main resonant series. The broad $2s2pnp$ $(^2D^e)$ and the narrow $2s2p3p(^2S^e)$ series dominate the total cross section. A much narrower $2s2pnf(^2D^e)$ series is visible near the second ionization threshold. As shown in Fig. 4(a), the boron spectrum computed with ASTRA exhibits a few additional ultra-narrow resonant features, highlighted by vertical tics (green online), that deserve some comments. By comparing the positions of

TABLE IV. Comparison of effective quantum number, n^* and resonance width Γ (a.u.), of $^2S^e$ resonances between $2s^2$ and $2s2p(^3P^o)$ thresholds.

Conf.	ASTRA n^*	ASTRA Γ	NEWSTOCK n^*	NEWSTOCK Γ	TAEC [111] n^*	TAEC [111] Γ
$2s2p3p$	2.765	4.99[–4]	2.743	5.22[–4]	2.733	4.70[–4]
$2s2p4p$	3.767	1.81[–4]	3.754	1.89[–4]	3.736	1.67[–4]
$2s2p5p$	4.766	8.51[–5]	4.756	8.96[–5]	4.736	7.75[–5]
$2s2p6p$	5.765	4.67[–5]	5.756	4.93[–5]	5.736	4.21[–5]
$2s2p7p$	6.765	3.15[–5]	6.756	2.99[–5]	6.735	2.54[–5]

TABLE V. Comparison of effective quantum number, n^* and resonance width Γ (a.u.), of $^2D^e$ resonances between $2s^2$ and $2s2p(^3P^o)$ thresholds.

Conf.	ASTRA		NEWSTOCK		TAEC [111]	
	n^*	Γ	n^*	Γ	n^*	Γ
$2s2p3p$	2.581	1.95[−3]	2.579	1.88[−3]	2.568	1.74[−3]
$2s2p4p$	3.596	6.75[−4]	3.594	6.46[−4]	3.585	5.68[−4]
$2s2p5p$	4.601	3.12[−4]	4.599	2.98[−4]	4.585	2.87[−4]
$2s2p6p$	5.601	1.70[−4]	5.601	1.62[−4]	5.588	1.54[−4]
$2s2p7p$	6.603	1.06[−4]	6.601	9.77[−5]	6.588	9.34[−5]
$2s2p4f$	4.013	1.56[−7]	4.013	1.07[−7]	4.009	1.57[−7]
$2s2p5f$	5.012	1.37[−7]	5.012	9.73[−8]	5.009	1.19[−7]
$2s2p6f$	6.012	1.32[−7]	6.011	7.27[−8]	6.008	8.65[−8]

the Rydberg bound states in NEWSTOCK and ASTRA, it can be shown that this series corresponds to $^2P^e$ bound states. These bound states manifest themselves as vanishingly narrow resonances in the photoionization spectrum due to the miniscule coupling of the $^2P^e$ bound states to the $^2D^e$ continuum. Similar minor symmetry mixing are to be expected in a molecular code because, in D_{2h} symmetry, the distinction between some $SO(3)$ irreducible representations is enforced dynamically, through numerical cancellations, rather than geometrically. That said, the mixing is small indeed. The largest of these resonances, at ~9.2 eV, has a width of just ≈10^{-13} a.u., comparable to machine precision, with the higher terms of the series being narrower still. Far from representing a failure of the code, therefore, such small values testify the remarkable accuracy with which ASTRA does reproduce the spherical symmetry of the atom.

The agreement between the two different *ab initio* approaches shows that the ASTRA formalism, based on non-standard high-order TDMs between ionic states with different multiplicities and designed for molecular systems, is consistent and can be used to obtain accurate results for atoms [119] as well.

B. Nitrogen molecule

The nitrogen molecule is an ideal system to test the accuracy of states in the ionization continuum, thanks to the availability of several theoretical and experimental benchmarks in the literature. In this subsection, we report on several comparisons, within the fixed-nuclei approximation (FNA). First, we ascertain that, to the minimal CIS level, ASTRA exactly reproduces the same result as an independent *ad hoc* CIS code, as it should. Next, we test ASTRA consistency and performance with the CC space constructed from correlated ions by comparing our results with those obtained with XCHEM [80], a state-of-the art molecular-ionization code, as well as with experimental values.

1. Comparison with ad hoc CIS Code

The ASTRA suite has been benchmarked against an *ad hoc* CIS code for the N_2 molecule, described in Appendix D. In the CIS model, the ionic states for the CC expansion are obtained by creating a vacancy in any of the valence orbitals
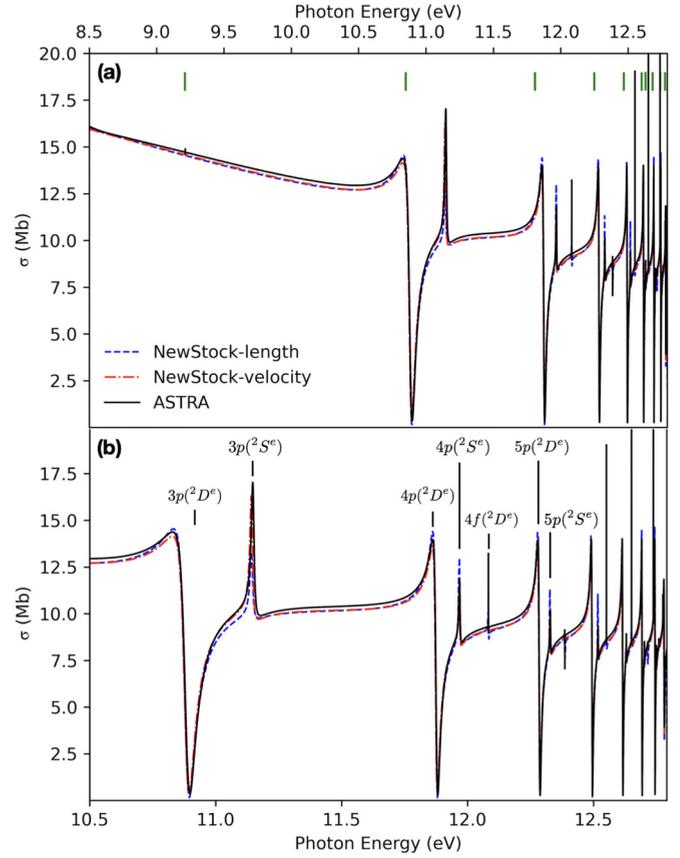
FIG. 4. (a) Total photoionization cross section between the $2s^2(^1S^e)$ and $2s2p(^3P^o)$ thresholds, from $^2P^o$ ground state of atomic boron, computed with ASTRA (solid black line) and compared with the length (dashed blue line) and the velocity (dash-dotted red line) gauge results from NEWSTOCK. The solid green lines indicate the positions of the bound states of $^2P^e$ symmetry. (b) Same cross section as in the top panel, magnified in the resonant region, which exhibits autoionizing states with both $^2S^e$ and $^2D^e$ symmetry. The notation $n\ell(^2L^e)$ denotes a resonance with dominant configuration $2s2pn\ell$ coupled to $^2L^e$ symmetry.

of the Hartree-Fock determinant of the N_2 ground state. All these single-configuration ions are subsequently coupled to all available VMOs and VHOs. The CC space obtained with this approach, therefore, coincides with the CIS space from the HF ground state of N_2. The HF ground state was computed using a 6-31G basis, whereas for the hybrid orbitals we used B splines of order six, with a node separation of 0.77 a.u. and angular momentum $\ell \leqslant 2$, within a 20 a.u. quantization box. The formulas for the transition-density matrices in the CIS basis, and the CC-matrix elements of the Hamiltonian and the dipole operators, are particularly simple, as shown in Appendix D. The density matrices were used to test, in the CIS case, the consistency of the TDMs generated by LUCIA, as well as that of the spin-adapted reduced TDMs evaluated by ASTRA. The matrix elements of the Hamiltonian, computed separately by the CIS code, were used to ascertain the correctness of the ASTRA structure code. The singlet states energy obtained from the Hamiltonian diagonalization in ASTRA agree with those computed with the CIS code within 10^{-12} a.u., which is

TABLE VI. Energy of and dipole oscillator strength (OS) between the ground state and the first few bright states of N_2, computed with XCHEM and ASTRA in the CI-singles limit, using a 6-31G basis for the HF orbitals. The energies are relative to ASTRA ground state.

State	Energy (eV)		OS (a.u.)	
	XCHEM	ASTRA	XCHEM	ASTRA
$X\ ^1\Sigma_g^+$	0.004	0.000		
$1\ ^1\Pi_u$	15.236	15.236	0.2062	0.2066
$1\ ^1\Sigma_u^+$	16.267	16.287	0.2053	0.2098
$2\ ^1\Pi_u$	16.311	16.321	0.0812	0.0802

compatible with the propagation of error entailed by the necessary operations, executed with machine precision.

2. Comparison with XCHEM in CI-singles subspace

ASTRA accounts for the matrix elements between polycentric GTOs and B splines, as well as for the exchange terms between hybrid and ionic orbitals. This feature allows us to use B splines across the whole radial range, thus providing the flexibility necessary to achieve the large photoelectron energies typical of core spectroscopies (several hundreds eV) as well as to describe the continuum of comparatively large molecules.

As detailed in the previous section, we ascertained that ASTRA predictions coincide with those of an *ad hoc* CIS code, when the same hybrid integrals are used. In the present section, we show that, in the CIS case, the results generated by ASTRA are compatible with those computed with XCHEM, an independent and well established suite of molecular photoionization codes. Like ASTRA, XCHEM is based on the CC *ansatz* and on the use of Gaussian-B-spline hybrid functions [79–82,84]. In contrast with ASTRA, however, XCHEM employs a different algorithm to evaluate CC-matrix elements, and it confines B splines beyond a certain distance R_0 to limit their overlap with the ionic polycentric GTOs. In XCHEM, the molecular and the external regions are bridged by a diffuse set of monocentric GTOs $G_{i\ell k}(r) = r^{\ell+2k} \exp(-\alpha_i r^2)$ [108].

Table VI compares the energy of the N_2 ground state and of the first few excited singlet states, as well as the dipole oscillator strengths from the former to the latter, obtained by diagonalizing the CC Hamiltonian computed either with XCHEM, using the 6-31G basis, or with ASTRA, using the 6-31G and the larger cc-pVQZ basis. In all these cases, the CC expansion includes the $X\ ^2\Sigma_g^+$, $A\ ^2\Pi_u$ and $B\ ^2\Sigma_u^+$ parent ions, augmented by single-electron states with $\ell \leqslant 2$, in a 400 a.u. quantization box. The ions were confined to a sphere with radius $R_{mol} = 20$ a.u. and two CAP were employed, starting at 100 and 200 a.u. with a strength of 10^{-6} and 10^{-5}, respectively. The ionic states are obtained from the HF configuration by removing one electron from the $3\sigma_g$, $1\pi_u$, and $2\sigma_u$ orbitals. The B splines were of order eight, the node separation chosen as $\Delta r = 0.5$ a.u. and for the diffuse GTOs in XCHEM we followed the same even-tempered set of exponents α_i defined in [81], with $k = 2$. The nitrogen molecule is aligned along the $\hat{z}$ axis and the internuclear distance is set to the experimental equilibrium value $R = 1.098$ Å [120]. The
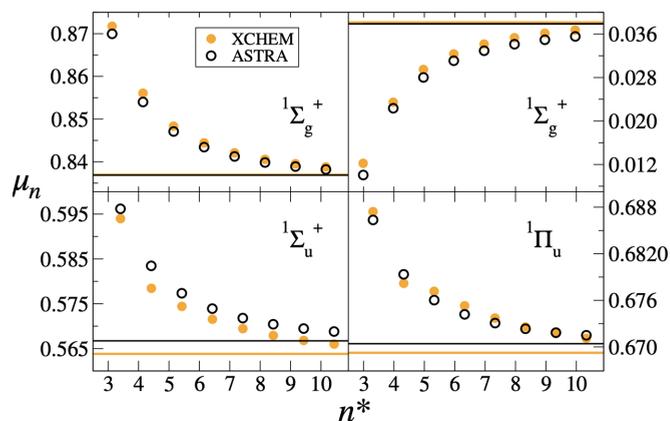
FIG. 5. Quantum defect μ_n^Γ vs the effective principal quantum number n^* for four Rydberg series of bound states in three symmetries, computed with XCHEM (filled circles, yellow online) and ASTRA (black empty circles) within the same CI-singles CC space used in Table VI. Horizontal lines indicate $\mu_\infty^\Gamma = \lim_{n\to\infty} \mu_n^\Gamma$.

energies computed with XCHEM and ASTRA using the 6-31G basis differ by no more than 0.02 eV, which is a remarkable agreement, given how different the virtual monoelectronic orbitals in the molecular region are for these two methods. The oscillator strength is more sensitive to the quality of the underlying electronic numerical basis. Nevertheless, ASTRA's results do not deviate more than 1% from those of XCHEM.

In Fig. 5, we compare the quantum defect $\mu_n \equiv n - n^*$ of four Rydberg series converging to the first ionization threshold $X\ ^2\Sigma_g^+$. The quantum defects obtained with the two methods are within 0.5% for all values of n^*. The agreement, therefore, is not restricted to the first few bound states, which may be well represented by GTOs alone, but it extends to Rydberg states with large quantum numbers and with a radial size well beyond the molecular region. These positive comparisons shows that, for the present calculation: (i) the two different configuration spaces made available to the Nth electron in XCHEM and ASTRA are equivalent, and (ii) ASTRA correctly employs the GBTOlib hybrid basic integrals.

3. CC with correlated ions

In this section, we go beyond the CIS approximation by considering the CC expansion obtained from the $X\ \Sigma_g^+$, $A\ ^2\Pi_u$, and $B\ ^2\Sigma_u^+$ correlated ions augmented by virtual orbitals with asymptotic angular momentum up to $\ell_{max} = 3$. The ions are computed by means of a CASCI calculation [92], over the $\{2\sigma_{g/u}, 3\sigma_{g/u}, 1\pi_u, 1\pi_g\}$ set of active orbitals, keeping the core $1\sigma_{g/u}$ orbitals inactive (doubly occupied), and employing three different GTO bases: 6-31G, cc-pVDZ, and cc-pVTZ. Table VII compares the energies of the first five singlet bound states obtained with ASTRA in the three different bases, alongside those obtained with XCHEM in the minimal 6-31G basis. All the energies are relative to the ground state obtained with ASTRA with the cc-pVTZ basis. For the XCHEM calculations, the same active space defined for ASTRA is employed to perform a CASCI and obtain the correlated ionic states. The same basis of diffused GTOs utilized in the CIS calculations is used to complement the polycentric GTOs at

TABLE VII. Energy of the first five singlet states of N_2, in eV, obtained by including correlated parent ions in the CC (see text for details) and using different GTO basis. The energy values are with respect to the ASTRA ground state when the basis set used to generate the molecular orbitals is cc-pVTZ.

State	XCHEM	ASTRA		
	6-31G	6-31G	cc-pVDZ	cc-pVTZ
$1\ ^1\Sigma_g^+$	0.639	0.631	0.128	0.000
$1\ ^1\Pi_g$	12.890	12.889	10.841	10.219
$1\ ^1\Sigma_u^-$	13.142	13.138	11.078	10.467
$2\ ^1\Sigma_g^+$	15.051	15.051	13.038	12.784
$1\ ^1\Delta_u$	15.636	15.639	13.622	13.411

short range. Once again, the agreement between XCHEM and ASTRA is excellent. In the present implementation, the CC expansion only consists of correlated ionic states augmented with an extra electron either in a bound or unbound orbital. A truncated CC expansion may be insufficient to accurately reproduce all the resonant states in the energy region of interest. This is because, in the CC expansion, the response of the ionic bound electrons to an additional electron is described only in terms of the expansion on a small set of states. While the CC captures well the polarization of the ion when the additional electron is located at several Bohr radii from the nucleus, it cannot describe well dynamic correlation at short range, such as the Coulomb hole [125], which has components in the multiple-ionization states of the ion itself. A standard way of dealing with this deficiency of the CC approach is to extend the expansion to include a pseudochannel with a large number of N-electron configurations built out of localized orbitals, without constraining $N-1$ electrons to any specific set of eigenstates of the ionic Hamiltonian. This pseudochannel, which is sometimes referred to as the $\mathcal{Q}$ space [126], can be very large. The CK-MESA code [76] has an efficient way of including the $\mathcal{Q}$ space via an optical potential built entirely out of GTOs [127]. XCHEM can also account for the $\mathcal{Q}$ space, even if it does so explicitly, which is more demanding than via an optical potential [76]. The inclusion of the $\mathcal{Q}$ space in ASTRA is outside the scope of the present paper. One viable approach to circumvent this limitation is extending the CC expansion so to include closed ionization channels whose parent ions are augmented only by AMO, VMO, and VHO, and in which the extra parent ions do not have to accurately reproduce any eigenstates of the ion.

We followed this approach to reproduce the high-resolution synchrotron radiation photoionization spectrum of N_2, in the vicinity of the Hopfield series of autoionizing states [81]. Besides the three parent ions $X\ ^2\Sigma_g^+$, $A\ ^2\Pi_u$, and $B\ ^2\Sigma_u^+$ we have used in the previous calculations, we added another 15 to the CC: $2\ ^2\Sigma_g^+$, $1\ ^2\Sigma_g^-$, $1\text{-}2\ ^2\Delta_g$, $1\ ^2\Phi_g$, $2\text{-}4\ ^2\Pi_u$, $2\ ^2\Sigma_u^+$, $1\text{-}2\ ^2\Delta_u$, $1\text{-}3\ ^2\Pi_g$, and $1\ ^2\Sigma_u^-$, to define the close ionization channels. In Fig. 6 we show the comparison of the photoionization cross section computed with ASTRA, the XCHEM result reported in [81] alongside four experimental spectra [121–124]. The theoretical signals have been convoluted with a 0.015 eV width normalized Gaussian to account for the experimental resolution. In the photoionization spectrum we can distin-
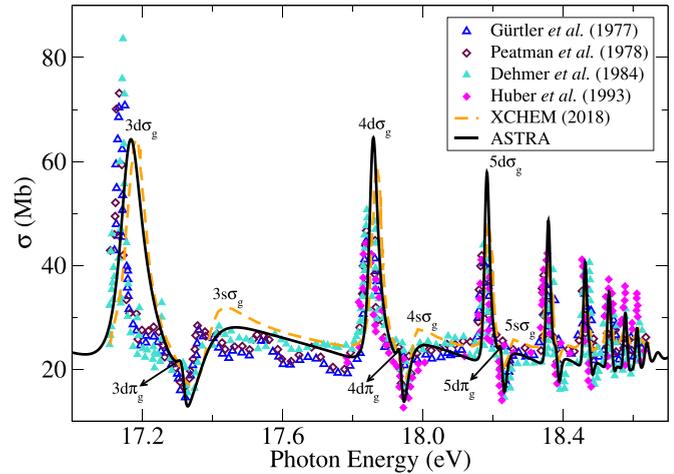
FIG. 6. Total photoionization cross section of N_2 in the Hopfield resonances region. The spectrum computed with ASTRA (solid black line) is compared with four independent measurements (triangles and diamonds) [121–124] and another theoretical spectrum obtained with XCHEM (dashed line) [81].

guish the three Hopfield series of resonances converging to the $B\ ^2\Sigma_u^+$ ionization threshold: $ns\sigma_g$, $nd\sigma_g$, and $nd\pi_g$, for which we indicate the first three terms. The ASTRA result is in very good agreement with both the XCHEM calculation and with the experimental spectra. For the first resonant feature corresponding to $3d\sigma_g$, both theoretical the results show a broader peak than the measured value, which is likely due to residual electronic correlation missing in the early terms of the Rydberg series.

The size of the close-coupling space, in this case, is 16 418 for the $A_g\ D_{2h}$ symmetry, and 10 918 for all the other D_{2h} symmetries considered. The calculations were conducted on a standalone workstation equipped with two 2.8 GHz Intel Xeon Platinum-8362 processors (32 cores each). While the current implementation of ASTRA is serial, LAPACK [128] diagonalization routines employ multithreaded drivers that occupied, on average, between 17 and 19 cores. The calculation of the CC matrix elements for all the operators, the full diagonalization of the real and complex Hamiltonian and the calculation of the photoionization cross section took ~2 hours.

C. Formaldehyde

To test ASTRA on a nonlinear polyatomic molecule, we selected formaldehyde, H_2CO, for which reasonably accurate experimental [129,130] and theoretical [70] photoionization cross sections are available. As for N_2, we used the FNA at the equilibrium geometry: $R_{CO} = 1.21$ Å, $R_{CH} = 1.12$ Å, $\angle HCH = 116.5°$ [131,132]. As detailed below, we conducted calculations with ASTRA at several levels of accuracy. In our largest calculation, the molecular orbitals are optimized by a state average MCSCF calculation on the first ten neutral state of the molecule, the VHO and ESO orbitals are computed with $\ell \leqslant 4$, and the CC expansion includes the first 37 molecular ions. The first ten ions are augmented with the full set of internal and external orbitals (AMO, VMO, VHO, and ESO), whereas the remaining 27 ions, which are all closed in
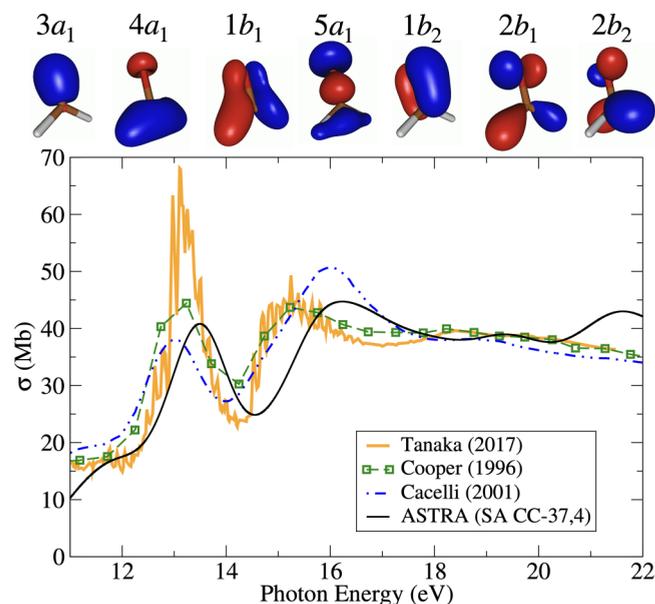
FIG. 7. Formaldehyde total photoionization cross section above the ionization threshold. The ASTRA result (black-solid line) is obtained by including 37 ions in the CC expansion with $\ell \leqslant 4$. The ions were obtained through a CASCI using the orbitals coming from a SA-MCSCF calculation (see text for details). The ASTRA spectrum is compared with the RPA calculation in [70] (dotted-dash line) and two measurements, with lower (dashed line and squares) and higher spectral resolution (orange-solid line) [129,130]. The top of the figure shows the first seven molecular orbitals included in the ASTRA active space, in ascending energy order, up to the LUMO ($2b_2$) [133,134].

the energy region of interest, are augmented by the internal orbitals only (AMO, VMO, and VHO).

Figure 7 compares the photoionization cross section of formaldehyde computed with ASTRA between 11 eV and 22 eV, using the highest level of electronic correlation described above, with Cacelli *et al.* RPA calculation [70], and with two measurements: one from a lower-resolution 1996 experiment by Cooper *et al.* [129] and the other from a higher-resolution 2017 experiment by Tanaka *et al.* [130]. The spectrum exhibits two prominent shape resonances, a narrower one near 13 eV, and a broader less pronounced one at 15 eV. In the higher-resolution experiment by Tanaka, the first peak has smaller width and larger height. Considering the different resolutions, however, the results of the two experiments are compatible with each other. The results from ASTRA are convoluted with a 1.2-eV Gaussian window, the same as in [70], in line with the ≈5% resolution reported by Cooper [129]. The RPA calculation [70] predicts well the position of the first peak, which it underestimates by just 0.2 eV, while it overestimates the position of the second peak by about 1 eV. Furthermore, the height of the first and second peak are significantly under- and over-estimated, respectively. These results suggest that both resonances are highly sensitive to correlation. In the ASTRA results, the position of the first peak is overestimated by about 0.2 eV, but its magnitude is in better agreement with the experiment by Cooper. Similarly to Cacelli's result, the position of the second peak predicted by ASTRA is about 1 eV
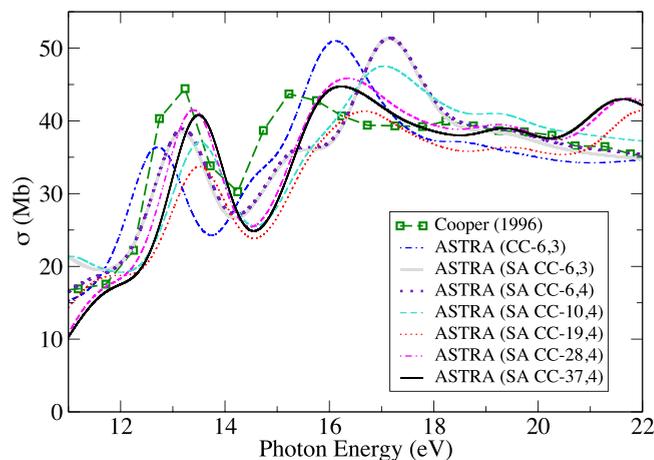
FIG. 8. Formaldehyde total photoionization cross section above the ionization threshold. Seven ASTRA spectra are shown together with Cooper's experiment ($\Delta E \approx 1.2$ eV) [129]. Six of the ASTRA calculations employed the same SA-MCSCF orbitals to perform the CASCI for the ionic states and they differ in the number of ions included in the CC expansion and in the maximum angular momentum considered for the partial waves. The format SA CC-N,L means that the calculation was performed by using the SA orbitals and a CC expansion with N ions and $\ell \leqslant L$.

above the experimental value. The peak height, however, is almost coincident with one observed in the experiment.

The results obtained with ASTRA are arguably the ones in best agreement with the experiment, to date. However, they also suggest that the theoretical model used here is still incomplete. Two possible causes for the discrepancy with the experiment are the residual dynamic correlation not captured by the CC expansion, and the fixed-nuclei approximation, which prevents us from accounting for the effects of nuclear motion and rearrangement in the molecular ion. On the electronic-structure side, ideally, the CC expansion should include a flexible and large pseudo-channel formed by arbitrary localized configurations. On the nuclear side, one should treat consistently nuclear motion and nuclear rearrangement on multiple excited electronic potential energy surfaces in a molecule with several degrees of freedom. While we plan to tackle both of these challenges in the future, these are formidable tasks beyond the scope of the present paper.

To illustrate the effect of the CC basis on formaldehyde photoionization cross-section, within the current capabilities of ASTRA, in Fig. 8 we compare the results obtained with seven different calculations, carried out with progressively more demanding parameters. In the first calculation, the CC expansion comprises six ionic states, $1\text{-}2\,^2B_1$, $1\,^2B_2$, $1\text{-}3\,^2A_1$, obtained from a CASCI calculation of 12 electrons in the nine active orbitals, $\{3a_1, 4a_1, 1b_1, 5a_1, 1b_2, 2b_1, 2b_2, 6a_1, 3b_1\}$, whereas the two core orbitals $1a_1$ and $2a_1$ (corresponding to the O1s and C1s K shells) were doubly occupied. These orbitals are obtained from the HF calculation of the neutral molecule within the cc-PVTZ GTO basis. To build the CC channels, each of these ions was augmented by an electron in any of the GTO (AMO, in the notation of Sec. III) and hybrid orbitals (VHO), and B splines with asymptotic angular momentum $\ell \leqslant 3$ spanning the whole quantization box (ESO). The B

spline and V_{CAP} parameters coincide with those used for nitrogen.

This first ASTRA calculation reproduces the characteristic double-hump structure of the spectrum and it is in very good agreement with the results by Cacelli *et al.* [70]. In particular, the absolute magnitude of the two peaks in the two independent calculations are remarkably similar. Both theoretical calculations, however, underestimate the position and magnitude of the first peak, and overestimate the energy and magnitude of the second peak, compared with the experiment.

In the other six ASTRA calculations, the molecular GTO orbitals are optimized through a SA-MCSCF step over the first 10 neutral states, with equal weights. The ions are subsequently determined by a CASCI calculation with the same number and type of inactive and active orbitals as in the first case.

In the second calculation, the CC space is generated from the first six ionic states only, augmented with orbitals with $\ell \leqslant 3$. The first and the second calculation, therefore, differ only in the optimization of the ion orbitals. The calculation reproduces very well the position of the first peak, quite possibly a coincidence, whereas the second peak splits into a shoulder around the position of the peak observed in the experiment and in a prominent third peak around 17 eV, with no correspondence in the measured data.

In the third calculation, the ions in the CC expansion coincide with those in the second calculation, but all the ions are now augmented with orbitals with $\ell \leqslant 4$. The second and the third calculation, therefore, differ only in the photoelectron partial-wave expansion. Both photoionization cross sections are almost identical except in the energy region approaching the ionization threshold ($E < 12$ eV), which indicates that in the region of interest, the partial-wave expansion is close to convergence for $\ell \leqslant 4$.

The fourth calculation coincides with the third, except that the CC expansion is extended to include the first 10 ions, each augmented by both the internal (AMO, VMO, and VHO) and by the external (ESO) partial-wave orbitals. The shoulder observed in the two previous spectra is greatly reduced and the magnitude of the second peak decreases, being closer to the experiment. The position of the first peak gets shifted 0.2 eV to higher energies and it's magnitude decreases by 2 Mb.

In the remaining three calculations, we investigated the convergence of the cross section using the same number of fully-augmented partial-wave channels as in the fourth, but adding 9, 18, and 27 other ions to the CC expansion, each augmented by the internal orbitals only. Since the corresponding channels are all closed in the energy region of interest, these last three calculations include progressively more dynamic correlation. Going from 10 to 19 to 28 ions, the position of the second peak is lowered by almost 1 eV. The position of the first peak is only marginally affected. The inclusion of nine other ions, from 28 to 37, has a comparatively small effect, which suggests that these calculations are close to convergence for the present choice of orbitals.

VI. CONCLUSIONS AND PERSPECTIVES

We have described a transition-density-matrix approach to the close-coupling method (TDMCC) for molecular pho-

toionization, implemented in the ASTRA code, and used it to compute the bound-state parameters, autoionizing-state parameters, and the single-photon integral photoionization cross section of benchmark atomic and molecular systems. The TDMCC method scales well with the size of the CI space of the parent ion, and it delivers results in excellent agreement with those in the literature.

In future studies, we will explore the scalability of AS-TRA by studying the ionization process in increasingly larger systems. Preliminary calculations indicate that ASTRA is capable of simulating the ionization of molecules as large as metallo-porphyrines, organometallic compounds comprised of 37 or more atoms, which opens a new way to the detailed predictions of wave-function-based methods for the ionization of molecules of biological interest.

From the methodological point of view, the ASTRA program has several natural directions along which to evolve. First, is the calculation of multichannel scattering states, which will allow us to determine photoelectron angular distributions [135,136], either in stationary regime, or as a result of the interaction of a target molecule with an ultrashort pulse of radiation. Second, is the extension of the calculation of the ionic states and of their TDMs from the present CASCI level to the more general RASCI [137–140] level. This extension will be particularly useful to describe highly conjugated systems. Third, we plan to use the SACC space as the basis for the time-dependent description of the photoionization wavepackets generated by the interaction of a target molecule with a sequence of ultrashort ionizing pulses [84,141,142]. This direction will allow us to reproduce transient-absorption spectra [143–146], four-wave mixing spectra [147–149], and photoelectron spectra [150] for state-of-the-art experiments.

In the longer term, we plan to explore also the many other natural extensions of the TDMCC formalism of AS-TRA, such as the inclusion of a Q space for the accurate description of short-range correlation [151], the inclusion of double-escape channels [152–154], to reproduce molecular double-ionization processes, as well as the description of multifragment systems [155–159] (e.g., a dissociating molecule, or a loosely bound aggregate) in terms of the TDMs of its separated components, in which the TDMs of the whole aggregate naturally factorizes, thus leading to drastic reduction in computational cost.

APPENDIX A: SECOND-QUANTIZATION FORMALISM

In this section we summarize the correspondence between antisymmetric electronic functions in the first-quantization coordinate representation and the associated Fock space, as well as some basic properties of many-body operators in

second quantization (SQ). For a general description of the SQ formalism, we refer the reader to the first chapter in [92].

Let us gather the spatial coordinate $\vec{r}$ and the spin coordinate $\zeta \in \{-\frac{1}{2}, \frac{1}{2}\}$ of an electron in the coordinate x = $(\vec{r}, \zeta)$. The state space for a single electron is spanned by a basis of spin orbitals $\phi_P(x) = \varphi_p(\vec{r})\,^2\chi_\pi(\zeta)$, where $\varphi_p(\vec{r}) \in \mathcal{L}^2(\mathbb{R}^3)$, $^2\chi_\pi(\zeta) = \langle \zeta | \pi \rangle = \delta_{\sigma\pi}$, and $P = (p, \pi)$ is a spin-orbital index, with $\pi \in \{-\frac{1}{2}, \frac{1}{2}\}$. In an N-electron system, the state space is spanned by Slater determinants $|\mathbf{P}\rangle$ built from arbitrary selections $\mathbf{P} = (P_1, \ldots, P_N)$ of N distinct spin orbitals from a complete one-particle basis, and defined as

$$\langle \mathbf{x} | \mathbf{P} \rangle = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_{P_1}(\mathbf{x}_1) & \phi_{P_2}(\mathbf{x}_1) & \cdots & \phi_{P_N}(\mathbf{x}_1) \\ \phi_{P_1}(\mathbf{x}_2) & \phi_{P_2}(\mathbf{x}_2) & \cdots & \phi_{P_N}(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{P_1}(\mathbf{x}_N) & \phi_{P_2}(\mathbf{x}_N) & \cdots & \phi_{P_N}(\mathbf{x}_N) \end{vmatrix}$$

$$= \sqrt{N!}\,\hat{\mathcal{A}}\,[\phi_{P_1}(\mathbf{x}_1)\phi_{P_2}(\mathbf{x}_2)\cdots\phi_{P_N}(\mathbf{x}_N)]$$
$$= |\phi_{P_1}\phi_{P_2}\cdots\phi_{P_N}|, \tag{A1}$$

where $\mathbf{x} = (x_1, x_2, \ldots, x_N)$ are the electronic coordinates, $\hat{\mathcal{A}} = \frac{1}{N!}\sum_{\mathcal{P} \in \mathcal{S}_N}(-1)^p\mathcal{P}$ is the antisymmetrizer, $\hat{\mathcal{A}}^2 = \hat{\mathcal{A}}$, $\hat{\mathcal{A}}^\dagger = \hat{\mathcal{A}}$, $\mathcal{S}_N$ is the symmetric group (group of permutations of N objects), and $\mathcal{P}$ and p are a permutation and its parity, respectively. These antisymmetric functions can be expressed in an equivalent way with the SQ formalism.

In SQ, the state of a system is expanded on a basis of occupation states that specify, for each spin orbital in an ordered list, if it is occupied by an electron or not. A state in which spin orbitals 1, 3, 4, and 7 are occupied, for example, is indicated by the ket $|1, 0, 1, 1, 0, 0, 1, 0, 0, 0, \ldots\rangle$. The total number of electrons in any such occupation state is given by $N_e = \sum_n n_i$, where n_i are the occupation numbers. A state with no electrons is represented by the so-called vacuum state, $|-\rangle = |0, 0, 0, 0, \ldots\rangle$, $\langle -|-\rangle = 1$. The space spanned by all occupation-number states with finite number of electrons is referred to as Fock space. To describe the action of arbitrary observables in this space, it is convenient to introduce operators that rise or lower the occupancy of a given spin orbital. The annihilation operator a_P removes an electron from the spin orbital P, if full, and annihilate the state otherwise. It acts on an occupation-number state as

$$a_P |k_1, \ldots, k_P, \ldots\rangle = \delta_{k_P 1}\,\Gamma_P^{\mathbf{k}}\,|k_1, \ldots, 0_P, \ldots\rangle, \tag{A2}$$

where $\Gamma_P^{\mathbf{k}} = (-1)^{k_1 + k_2 + \cdots + k_{P-1}}$. The adjoint of a_P, $a_P^\dagger$, adds an electron to spin orbital P, if empty, and annihilate the state otherwise,

$$a_P^\dagger |k_1, \ldots, k_P, \ldots\rangle = \delta_{k_P 0}\,\Gamma_P^{\mathbf{k}}\,|k_1, \ldots, 1_P, \ldots\rangle. \tag{A3}$$

Creation and annihilation operators satisfy the following well-known anticommutation relations,

$$\{a_I, a_J\} = \{a_I^\dagger, a_J^\dagger\} = 0, \quad \{a_I^\dagger, a_J\} = \delta_{IJ}. \tag{A4}$$

The correspondence between first- and second-quantization formalism is completed by a phase convention for the equivalence between Slater determinants and occupation-number states. Here, we adopt the following convention:

$$|\phi_{P_1}\phi_{P_2}\cdots\phi_{P_N}| \leftrightarrow a_{P_1}^\dagger a_{P_2}^\dagger \cdots a_{P_N}^\dagger |-\rangle. \tag{A5}$$

a. Operators in second-quantization formalism

In first quantization, the definition of one-body operators, $\langle \mathbf{x}|\hat{O}|\mathbf{x}'\rangle = \sum_{i=1}^{N_e} o(x_i; x_i')$, and of the two-body Coulomb repulsion operator, $\langle \mathbf{x}, \mathbf{x}'|\hat{G}|\mathbf{x}, \mathbf{x}'\rangle = \sum_{i,j<i} |\vec{r}_i - \vec{r}_j'|^{-1}$, depend explicitly on the number of particles in the system. In SQ, on the other hand, these operators do not explicitly depend on the number of electrons, which is a major advantage,

$$\hat{O} = \sum_{RS} o_{RS} a_R^\dagger a_S, \tag{A6}$$

$$\hat{G} = \frac{1}{2}\sum_{PQRS}[PQ|RS]\,a_P^\dagger a_R^\dagger a_S a_Q, \tag{A7}$$

where

$$o_{RS} = \iint dx\,dx'\,\phi_R^*(x)o(x, x')\phi_S(x'),$$

$$[PQ|RS] = \iint dx\,dx'\,\frac{\phi_P^*(x)\phi_Q(x)\,\phi_R^*(x')\phi_S(x')}{r_{12}}.$$

In ASTRA, we need to evaluate matrix elements of strings of creation and annihilation operators between states $|A\rangle$ and $|B\rangle$ with well-defined spin, S_A, S_B, and spin projections, Σ_A, Σ_B, respectively. To this end, it is convenient to expand the operators in spherical tensors $\mathcal{O}_{T\tau}$, and make use of Wigner-Eckart theorem,

$$\langle A_{\Sigma_A}|\mathcal{O}_{T\tau}|B_{\Sigma_B}\rangle = \frac{C_{S_B\Sigma_B, T\tau}^{S_A\Sigma_A}}{\sqrt{2S_A + 1}}\,\langle A\|\mathcal{O}_T\|B\rangle, \tag{A8}$$

where $C_{a\alpha, b\beta}^{c\gamma}$ are Clebsch-Gordan coefficients and $\langle A\|\mathcal{O}_T\|B\rangle$ is a reduced matrix element [95]. To treat spin tensors $\mathcal{O}_{T\tau}$ consistently, it is necessary to adopt the convention on the phase imparted to these operators by the ladder spin operators,

$$[\hat{S}_\pm, \mathcal{O}_{T\tau}] = \sqrt{T(T+1) - \tau(\tau \pm 1)}\,\mathcal{O}_{T\tau\pm1},$$
$$[\hat{S}_z, \mathcal{O}_{T\tau}] = \tau\,\mathcal{O}_{T\tau}, \tag{A9}$$

where the rising and lowering operators $\hat{S}_\pm = \hat{S}_x \pm i\hat{S}_y$ are related to the spherical tensor component of spin through $\hat{S}_\pm = \mp\sqrt{2}S_{\pm1}$. The convention in (A9) is consistent with [160] [see Eq. (2.3.1)], as well as with [95] [see Eq. (1) in ch. 3].

For each orbital p, the two creation operators $a_{p_\pi}^\dagger$, with $\pi = \pm\frac{1}{2}$, comply with the phase conventions (A9), whereas the adjoint operators $a_{p_\pi} \equiv (a_{p_\pi}^\dagger)^\dagger$ do not. For this reason, it is convenient to define a second set of tensors operators $b_{p_\pi}^\dagger$ and b_{p_π},

$$b_{p_\pi}^\dagger = a_{p_\pi}^\dagger, \quad b_{p_\pi} = (-1)^{\frac{1}{2}+\pi}a_{p_{-\pi}}, \tag{A10}$$

which satisfy (A9), as well as the convention

$$(b_{p_\pi})^\dagger = (-1)^{\frac{1}{2}+\pi}b_{p_{-\pi}}^\dagger. \tag{A11}$$

Two spherical tensors $A_{T\tau}$ and $B_{J\mu}$ can be coupled to a rank-K tensor as

$$[A_T \otimes B_J]_{K\kappa} \equiv \sum_{\tau\mu}C_{T\tau, J\mu}^{K\kappa}A_{T\tau}B_{J\mu}. \tag{A12}$$

Conversely, the product of two spherical tensors can be expanded in spin-coupled products

$$A_{T\tau}B_{J\mu} = \sum_{K\kappa} C^{K\kappa}_{T\tau,J\mu}[A_T \otimes B_J]_{K\kappa}. \qquad \text{(A13)}$$

In particular,

$$a_P^\dagger a_Q^\dagger = a_{p_\pi}^\dagger a_{q_\theta}^\dagger = b_{p_\pi}^\dagger b_{q_\theta}^\dagger = \sum_{T\tau} C^{T\tau}_{\frac{1}{2}\pi,\frac{1}{2}\theta}[b_p^\dagger \otimes b_q^\dagger]_{T\tau},$$

$$a_P^\dagger a_Q^\dagger a_R^\dagger = \sum_{J\mu T\tau} C^{J\mu}_{\frac{1}{2}\pi,T\tau} C^{T\tau}_{\frac{1}{2}\theta,\frac{1}{2}\rho}[b_p^\dagger \otimes [b_q^\dagger \otimes b_r^\dagger]_T]_{J\mu}. \quad \text{(A14)}$$

These expressions, when used in combination with Wigner-Eckart's theorem, allow us to compute matrix elements for arbitrary values of n spin magnetic quantum numbers from selected matrix elements in which two of those n quantum numbers are fixed,

$$\langle A\|\mathcal{O}_T\|B\rangle = \frac{\sqrt{2S_A+1}}{C^{S_A M}_{S_B M,T0}}\, \langle A_M|\mathcal{O}_{T0}|B_M\rangle. \qquad \text{(A15)}$$

This step is essential for the current ASTRA implementation, which relies on quantum-chemistry calculations that consider a single well-defined spin-projection M at a time. For tensors with an equal number of creation and annihilation operators, with $T = 0, 1, 2$, it is always possible to find an M such that $C^{S_B M}_{S_A M,T0} \neq 0$. Notice that

$$[b_p \otimes b_q]_{T\tau} = (-1)^{1-T}[b_q \otimes b_p]_{T\tau},$$
$$[b_p^\dagger \otimes b_q^\dagger]_{T\tau} = (-1)^{1-T}[b_q^\dagger \otimes b_p^\dagger]_{T\tau}. \qquad \text{(A16)}$$

APPENDIX B: TRANSITION-DENSITY MATRICES

ASTRA evaluates matrix elements within a CC space obtained by augmenting a given set of multiconfiguration ionic states with arbitrary single-electron functions. To this end, it is convenient to define the first-, second-, and third-order transition-density matrices (TDMs) as the matrix elements between ionic states of strings containing one, two, or three creator operators followed by an equal number of annihilation operators (normal order),

$$\rho^{BA}_{QP} = \langle A|a_P^\dagger a_Q|B\rangle, \qquad \text{(B1)}$$

$$\pi^{BA}_{RS,PQ} = \langle A|a_P^\dagger a_Q^\dagger a_S a_R|B\rangle, \qquad \text{(B2)}$$

$$\gamma^{BA}_{STU,PQR} = \langle A|a_P^\dagger a_Q^\dagger a_R^\dagger a_U a_T a_S|B\rangle. \qquad \text{(B3)}$$

The rest of this section describes the coupling scheme used in ASTRA to express these TDMs in terms of their reduced counterpart.

b. Inactive, active, and virtual orbitals.

In the expressions that involve TDMs between ionic states, it is convenient to distinguish between: (i) inactive spin orbitals, which are identically represented in all of the ions, and which we designate with the letters W, X, Y, Z; (ii) active orbitals, i.e., those without a well defined occupation in the CI of the ion; and (iii) virtual orbitals, which lie outside the active space, and hence have zero occupation in all the ions. Any TDM element with virtual-orbitals indexes, of course, is

zero. The matrix elements with inactive orbitals, which also often vanish, admit simple expressions in terms of lower-order density matrices. In particular,

$$\rho^{BA}_{QX} = \rho^{BA}_{XQ} = \delta_{AB}\delta_{XQ},$$

$$\pi^{BA}_{XS,PQ} = \pi^{BA}_{SX,QP}$$
$$= (1 - \delta_{XS})(1 - \delta_{PQ})(\delta_{PX}\rho^{BA}_{SQ} - \delta_{QX}\rho^{BA}_{SP}),$$

$$\pi^{BA}_{RS,PX} = \pi^{BA}_{SR,XP}$$
$$= (1 - \delta_{RS})(1 - \delta_{PX})(\delta_{SX}\rho^{BA}_{RP} - \delta_{RX}\rho^{BA}_{SP}),$$

$$\pi^{BA}_{RS,XY} = \pi^{BA}_{XY,SR}$$
$$= \delta_{AB}(1 - \delta_{RS})(1 - \delta_{XY})(\delta_{SY}\delta_{RX} - \delta_{SX}\delta_{RY}),$$

$$\pi^{BA}_{XP,YQ} = -\pi^{BA}_{XP,QY}$$
$$= (1 - \delta_{XP})(1 - \delta_{YQ})(\delta_{XY}\rho^{BA}_{PQ} - \delta_{QX}\delta_{PY}\delta_{AB}),$$

$$\pi^{BA}_{XY,ZQ} = \delta_{AB}(1 - \delta_{XY})(1 - \delta_{ZQ})(\delta_{XZ}\delta_{QY} - \delta_{QX}\delta_{YZ}).$$

These formulas allow us to incorporate the two-body interactions with the electrons in the core orbitals in effective one-body operators.

c. Reduced transition-density matrices

The quantum chemistry code LUCIA does not provide arbitrary TDM elements, since it restricts the total spin projection to be the same for the two ions. However, within this constraint, LUCIA does provide a sufficient number of *independent* matrix elements to reconstruct all the others by means of the Wigner Eckart theorem (WE). To take advantage of this theorem, it is necessary to cast the string of operators involved in the TDM in terms of well-defined spin tensors. As discussed in the previous section, the first step to use creators and destructor operators consistently with the phase conventions for spherical tensors, we have to express them in terms of the b and $b^\dagger$ operators. Subsequently, we can apply WE to express any TDM elements in terms of reduced matrix elements of its tensorial components.

In the case of the one-body TDM,

$$\rho^{BA}_{QP} = (-1)^{\frac{1}{2}-\theta} \langle A|b_{p\pi}^\dagger b_{q-\theta}|B\rangle$$

$$= (-1)^{\frac{1}{2}-\theta} \sum_{T\tau} C^{T\tau}_{\frac{1}{2}\pi,\frac{1}{2}-\theta} \langle A|[b_p^\dagger \otimes b_q]_{T\tau}|B\rangle$$

$$= \frac{(-1)^{\frac{1}{2}-\theta}}{\Pi_{S_A}} \sum_T \mathsf{R}^{BA}_{[q,p]_T} \sum_\tau C^{T\tau}_{\frac{1}{2}\pi,\frac{1}{2}-\theta} C^{S_A \Sigma_A}_{S_B \Sigma_B, T\tau} \quad \text{(B4)}$$

where $\Pi_{ab\cdots} = \sqrt{(2a+1)(2b+1)\cdots}$ and we have introduced the notation $\mathsf{R}^{BA}_{[q,p]_T}$ to designate the reduced one-body TDM,

$$\mathsf{R}^{BA}_{[q,p]_T} = \langle A\|[b_p^\dagger \otimes b_q]_T\|B\rangle. \qquad \text{(B5)}$$

To evaluate the reduced one-body TDMs from selected matrix elements between states with the same spin projection, let us consider the cases $T = 0, 1$ separately. For $T = 0$, the matrix $\mathsf{R}^{BA}_{[q,p]_0}$ is nonzero only if the two ions have the same multiplicity. For any value of their projection $\Sigma_A = \Sigma_B$, we

can compute the reduced matrix element as

$$R^{BA}_{[q,p]_0} = \sqrt{S_A + 1/2} \sum_\sigma \rho^{BA}_{q_\sigma p_\sigma}. \tag{B6}$$

For $T = 1$, the reduced matrix element is zero if both A and B are singlet states. For parent ions with an odd number of electrons, it is possible to compute $R^{BA}_{[q,p]_1}$ from the uncoupled density matrix between the states with $\Sigma_A = \Sigma_B = \Sigma = 1/2$,

$$R^{BA}_{[q,p]_1} = \frac{\sqrt{S_A + 1/2}}{C^{S_A 1/2}_{S_B 1/2, 10}} \left(\rho^{BA}_{q_\alpha p_\alpha} - \rho^{BA}_{q_\beta p_\beta} \right). \tag{B7}$$

For parent ions with an even number of electrons, the matrix element between ions with different spin can be computed choosing $\Sigma_A = \Sigma_B = \Sigma = 0$ (which is required if one of the two spin is zero). For parent ions with the same, nonzero spin, however, it is necessary to use any other spin projection (e.g., $\Sigma_A = \Sigma_B = \Sigma = 1$)

$$R^{BA}_{[q,p]_1} = \frac{\sqrt{S_A + 1/2}}{C^{S_A \Sigma}_{S_B \Sigma, 10}} \left(\rho^{BA}_{q_\alpha p_\alpha} - \rho^{BA}_{q_\beta p_\beta} \right). \tag{B8}$$

In the two-body TDM, we first couple the spin of the two creation operators and those of the two annihilation operators to a total spin $T = 0$, 1 and $J = 0$, 1, respectively. These are subsequently coupled to a total spin $K = 0$, 1, 2,

$$\pi^{BA}_{RS,PQ} = -(-1)^{-\sigma-\rho} \langle A | b^\dagger_{p\pi} b^\dagger_{q\theta} b_{s-\sigma} b_{r-\rho} | B \rangle$$
$$= -(-1)^{-\sigma-\rho} \Pi^{-1}_{S_A} \sum_{JTK} \Pi^{BA}_{[[sr]_J,[pq]_T]_K}$$
$$\times \sum_{\mu\tau\kappa} C^{S_A \Sigma_A}_{S_B \Sigma_B, K\kappa} C^{K\kappa}_{T\tau, J-\mu} C^{J-\mu}_{\frac{1}{2}-\sigma, \frac{1}{2}-\rho} C^{T\tau}_{\frac{1}{2}\pi, \frac{1}{2}\theta} \tag{B9}$$

where the reduced matrix element $\Pi^{BA}_{[[sr]_J,[pq]_T]_K}$ is

$$\Pi^{BA}_{[[sr]_J,[pq]_T]_K} = \langle A \| [[b^\dagger_p \otimes b^\dagger_q]_T \otimes [b_s \otimes b_r]_J]_K \| B \rangle.$$

As shown in (A16), spin-coupled electron pairs anticommute or commute depending on whether they form a singlet or a triplet, respectively, which implies the following permutation symmetries:

$$\Pi^{BA}_{[[sr]_J,[pq]_T]_K} = (-1)^{J+1} \Pi^{BA}_{[[rs]_J,[pq]_T]_K}$$
$$= (-1)^{T+1} \Pi^{BA}_{[[sr]_J,[qp]_T]_K}$$
$$= (-1)^{J+T} \Pi^{BA}_{[[rs]_J,[qp]_T]_K}. \tag{B10}$$

The two-body reduced matrix elements have the following expression in terms of TDMs between states with the same spin projection $\Sigma_A = \Sigma_B = \Sigma$,

$$\Pi^{BA}_{[[sr]_J,[pq]_T]_K} = \frac{(-1)^J \Pi_{S_A}}{C^{S_A \Sigma}_{S_B \Sigma, K0}} \sum_{\pi\theta\sigma\rho\tau} (-1)^\tau$$
$$\times C^{K0}_{T\tau, J-\tau} C^{J\tau}_{\frac{1}{2}\sigma, \frac{1}{2}\rho} C^{T\tau}_{\frac{1}{2}\pi, \frac{1}{2}\theta} \pi^{BA}_{r_\rho s_\sigma, p_\pi q_\theta}.$$

Finally, the coupling scheme for the 3B-TDM is

$$\gamma^{BA}_{STU,PQR} = \zeta \langle A | b^\dagger_{p\pi} b^\dagger_{q\theta} b^\dagger_{r\rho} b_{u-\upsilon} b_{t-\tau} b_{s-\sigma} | B \rangle$$
$$= \frac{\zeta}{\Pi_{S_A}} \sum_{CDFJK} \Gamma^{BA}_{[[[ut]_F s]_J, [[pq]_C r]_D]_K}$$

$$\times \sum_{\gamma\delta\varphi\mu\kappa} C^{C\gamma}_{\frac{1}{2}\pi, \frac{1}{2}\theta} C^{D\delta}_{C\gamma, \frac{1}{2}\rho} C^{F\varphi}_{\frac{1}{2}-\upsilon, \frac{1}{2}-\tau}$$
$$\times C^{J\mu}_{F\varphi, \frac{1}{2}-\sigma} C^{K\kappa}_{D\delta, J\mu} C^{S_A \Sigma_A}_{S_B \Sigma_B, K\kappa} \tag{B11}$$

where $\zeta = (-1)^{-1/2-\upsilon-\tau-\sigma}$ and

$$\Gamma^{BA}_{[[[ut]_F s]_J, [[pq]_C r]_D]_K}$$
$$= \langle A \| [[[b^\dagger_p \otimes b^\dagger_q]_C \otimes b^\dagger_r]_D \otimes [[b_u \otimes b_t]_F \otimes b_s]_J]_K \| B \rangle. \tag{B12}$$

APPENDIX C: MATRIX ELEMENTS BETWEEN CC STATES

This Appendix derives the overlap as well as the matrix elements of one-body and two-body operators between CC states, without and with spin adaptation. The evaluation of each of the relevant matrix elements between CC states starts with the reduction of the corresponding operator string.

1. Overlap

The CC states are generally not orthogonal, so it is necessary to evaluate their overlap,

$$\langle A, P | B, Q \rangle = \langle A | a_P a^\dagger_Q | B \rangle. \tag{C1}$$

In normal form, the string of operators is

$$a_P a^\dagger_Q = S_{PQ} - a^\dagger_Q a_P, \tag{C2}$$

where $S_{PQ} = \langle P | Q \rangle$, and hence

$$\langle A, P | B, Q \rangle = \langle A | a_P a^\dagger_Q | B \rangle = S_{PQ} \delta_{AB} - \rho^{BA}_{PQ}. \tag{C3}$$

The CC states built from internal active orbitals, therefore, in general are not normalized,

$$\langle A, P | A, P \rangle = 1 - \rho^{AA}_{PP}. \tag{C4}$$

Using (C3) and (11), we obtain the expression for the overlap between SACC states,

$$\langle A, p | B, q \rangle = \sum_{\Sigma_A \Sigma_B \pi\theta} C^{S\Sigma}_{S_A \Sigma_A, \frac{1}{2}\pi} C^{S\Sigma}_{S_B \Sigma_B, \frac{1}{2}\theta} \left(S_{pq} \delta_{\pi,\theta} \delta_{AB} - \rho^{BA}_{p_\pi q_\theta} \right)$$
$$= \delta_{AB} S_{pq} - \sum_{\Sigma_A \Sigma_B \pi\theta} C^{S\Sigma}_{S_A \Sigma_A, \frac{1}{2}\pi} C^{S\Sigma}_{S_B \Sigma_B, \frac{1}{2}\theta} \rho^{BA}_{p_\pi q_\theta}.$$

Summing the Clebsch-Gordan coefficients, we can rewrite the overlap as

$$\langle A, p | B, q \rangle = \delta_{AB} S_{pq} + W^{BA}_{pq},$$

where

$$W^{BA}_{pq} = \eta \sum_T R^{BA}_{[p,q]_T} \Pi_T \begin{Bmatrix} T & 1/2 & 1/2 \\ S & S_B & S_A \end{Bmatrix},$$

and $\eta = (-1)^{S_B - 1/2 + S}$.

2. One-body operators

The matrix element between CC states of a generic spin-free one-body operator is

$$\langle A, P | \hat{O} | B, Q \rangle = o_{RS} \langle A | a_P a^\dagger_R a_S a^\dagger_Q | B \rangle, \tag{C5}$$

where summation over repeated magnetic quantum numbers and orbital indexes is implied. In normal form, the operator string reads

$$a_P a_R^\dagger a_S a_Q^\dagger = \delta_{PR}\delta_{QS} + S_{PQ} a_R^\dagger a_S - \delta_{PR} a_Q^\dagger a_S$$
$$- \delta_{QS} a_R^\dagger a_P + a_Q^\dagger a_R^\dagger a_P a_S, \qquad (C6)$$

and hence the matrix element becomes

$$\langle A, P|\hat{O}|B, Q\rangle = o_{PQ}\,\delta_{AB} + S_{PQ}\,O_{AB}$$
$$- o_{PS}\,\rho_{S,Q}^{BA} - \rho_{P,R}^{BA}\,o_{RQ} + o_{RS}\,\pi_{SP,QR}^{BA}$$

where $O_{AB} = \langle A|\hat{O}|B\rangle$. Notice that we can evaluate O_{AB} within ASTRA starting from the one-body TDM provided by LUCIA. Such one-body spin-free operator vanish unless $\Sigma_A = \Sigma_B$,

$$\langle A|\hat{O}|B\rangle = o_{rs}\langle A|a_{r\sigma}^\dagger a_{s\sigma}|B\rangle = o_{rs}\rho_{s_\sigma r_\sigma}^{BA}. \qquad (C7)$$

Since the two spin projections coincide, the expression for ρ becomes

$$\rho_{s_\sigma r_\sigma}^{BA} = \sum_T \frac{\mathsf{R}_{[s,r]_T}^{BA}}{\Pi_{S_A}} \sum_{\tau\sigma} (-1)^{\frac{1}{2}-\sigma} C_{\frac{1}{2}\sigma,\frac{1}{2}-\sigma}^{T\tau} C_{S\Sigma,T\tau}^{S\Sigma}$$
$$= \sum_T \frac{C_{S\Sigma,T0}^{S\Sigma}\mathsf{R}_{[s,r]_T}^{BA}}{\Pi_{S_A}} \sum_\sigma (-1)^{\frac{1}{2}-\sigma} C_{\frac{1}{2}\sigma,\frac{1}{2}-\sigma}^{T0}$$
$$= \frac{\mathsf{R}_{[s,r]_0}^{BA}}{\sqrt{2}\,\Pi_{S_A}}.$$

We introduce here the quantity

$$Q_{sr}^{BA} = \frac{\sqrt{2}\,\mathsf{R}_{[s,r]_0}^{BA}}{\Pi_{S_A}} \qquad (C8)$$

with which

$$\langle A|\hat{O}|B\rangle = o_{rs}\frac{\Pi_{S_A}^{-1}}{\sqrt{2}}\mathsf{R}_{[s,r]_0}^{BA} = o_{rs}Q_{sr}^{BA}/2. \qquad (C9)$$

The extra factor of 2 in the definition of Q_{rs}^{BA} is the legacy of an arbitrary convention. As shown further down, Q_{rs}^{BA} is contained also in the expression of two-body operators, where it appears without prefactors.

It is now possible to obtain the expression for the matrix element between SACC states. After some straightforward recoupling of tensor operators and contraction of Clebsch-Gordan coefficients, one finds

$$\langle A, p|\hat{O}|B, q\rangle$$
$$= S_{pq}\,O_{AB} + o_{pq}\,\delta_{AB}$$
$$+ \eta \sum_T \left(\mathsf{R}_{[p,r]_T}^{BA}\,o_{rq} + o_{ps}\mathsf{R}_{[s,q]_T}^{BA}\right)\Pi_T \begin{Bmatrix} T & 1/2 & 1/2 \\ S & S_B & S_A \end{Bmatrix}$$
$$+ \eta\,o_{rs}\sum_{KTJ}(-1)^{J+K}\Pi_{[[ps]_J,[qr]_T]_K}^{BA}\,\Pi_{KTJ}$$
$$\times \begin{Bmatrix} S_A & K & S_B \\ 1/2 & S & 1/2 \end{Bmatrix}\begin{Bmatrix} J & T & K \\ 1/2 & 1/2 & 1/2 \end{Bmatrix}. \qquad (C10)$$

We can condense the notation by using the tensor W_{pq}^{BA} introduced above, and the new tensor

$$P_{ps,qr}^{BA} = \eta\sum_{KTJ}(-1)^{J+K}\Pi_{[[ps]_J,[qr]_T]_K}^{BA}\,\Pi_{KTJ}$$
$$\times \begin{Bmatrix} S_A & K & S_B \\ 1/2 & S & 1/2 \end{Bmatrix}\begin{Bmatrix} J & T & K \\ 1/2 & 1/2 & 1/2 \end{Bmatrix}, \qquad (C11)$$

with which the matrix element becomes

$$\langle A, p|\hat{O}|B, q\rangle = S_{pq}\,O_{AB} + o_{pq}\,\delta_{AB}$$
$$+ W_{pr}^{BA}\,o_{rq} + o_{ps}\,W_{sq}^{BA} + o_{rs}\,P_{ps,qr}^{BA}. \qquad (C12)$$

The orbitals used to augment the ions do not include inactive orbitals. As a consequence, the summation over the orbital indexes in the third and fourth term can be restricted to the active orbitals only. On the other hand, inactive orbitals do give a contribution to the fifth term, which can be expressed in terms of one-body density matrices. By explicitly indicating the summation over orbitals indexes, and labeling with a prime the sums over active orbitals only, the result reads

$$\langle A, p|\hat{O}|B, q\rangle = S_{pq}\,O_{AB} + o_{pq}\,\delta_{AB} + \langle O\rangle_{\text{core}}\,W_{pq}^{BA}$$
$$+ \sideset{}{'}\sum_r \left(W_{pr}^{BA}\,o_{rq} + o_{pr}W_{rq}^{BA}\right) + \sideset{}{'}\sum_{rs} o_{rs}P_{ps,qr}^{BA}, \qquad (C13)$$

where $\langle O\rangle_{\text{core}} = 2\sum_x o_{xx}$. Notice that only when both ions are augmented by an active internal orbital do all terms in the expression contribute to the matrix element. If one of the two orbitals is external, only one of the terms involving the B tensors survive. If both orbitals are external, only the first two terms survive.

3. Two-body operators

The matrix element of a two-body spin-free operator between CC states is

$$\langle A, P|\hat{G}|B, Q\rangle = \tfrac{1}{2}[TU|RS]\langle A|a_P a_T^\dagger a_R^\dagger a_S a_U a_Q^\dagger|B\rangle. \qquad (C14)$$

A term on the right-hand side vanish, of course, unless the spin projections of T and U, and of R and S coincide. For the sake of clarity, however, we do not split the orbital and spin indexes yet. In normal form, the operator string reads

$$a_P a_T^\dagger a_R^\dagger a_S a_U a_Q^\dagger$$
$$= \delta_{PR}\delta_{QS}\,a_T^\dagger a_U - \delta_{UQ}\delta_{PR}\,a_T^\dagger a_S$$
$$- \delta_{TP}\delta_{QS}\,a_R^\dagger a_U + \delta_{TP}\delta_{UQ}\,a_R^\dagger a_S$$
$$+ \delta_{PR}\,a_T^\dagger a_Q^\dagger a_U a_S + \delta_{QS}\,a_T^\dagger a_R^\dagger a_U a_P$$
$$- \delta_{PQ}\,a_T^\dagger a_R^\dagger a_U a_S + \delta_{UQ}\,a_T^\dagger a_R^\dagger a_P a_S$$
$$+ \delta_{TP}\,a_Q^\dagger a_R^\dagger a_U a_S + a_T^\dagger a_Q^\dagger a_R^\dagger a_U a_P a_S. \qquad (C15)$$

The general expression for the matrix element, considering that the operator is spin free, becomes

$$\langle A, P|\hat{G}|B, Q\rangle$$
$$= G_{AB}S_{PQ} + \delta_{\pi\theta}[pq|rs]\rho_{s_\rho,r_\rho}^{BA} - [ps|rq]\rho_{s_\pi,r_\theta}^{BA}$$

$$+ [pt|rs]\pi^{BA}_{t_\pi s_\rho, r_\rho q_\theta} + [qt|rs]\pi^{BA}_{p_\pi s_\rho, r_\rho t_\theta}$$
$$- \tfrac{1}{2}[tu|rs]\gamma^{BA}_{u_\tau s_\rho p_\pi, t_\tau r_\rho q_\theta}. \tag{C16}$$

We can now obtain an expression for the matrix element of the two-body operator between SACC states. The full expression of the result is not only lengthy to derive but also rather long. Luckily, as in the case of the overlap and of the one-body operators, it is possible to express it in condensed form with the help of *ad hoc* auxiliary tensors. We have already defined three of these tensors, Q^{BA}_{sr}, W^{BA}_{rs}, and $P^{BA}_{rs,tu}$. It is now convenient to define

$$C^{BA}_{psu,trq} = \frac{\eta}{2} \sum_{CDFJK} (-1)^{K+D+C-1/2} \Pi_{CDFJK}$$
$$\times \begin{Bmatrix} S_A & K & S_B \\ 1/2 & S & 1/2 \end{Bmatrix} \begin{Bmatrix} J & C & 1/2 \\ 1/2 & K & D \end{Bmatrix} \begin{Bmatrix} C & 1/2 & J \\ F & 1/2 & 1/2 \end{Bmatrix}$$
$$\times \Gamma^{BA}_{[[[ps]_F u]_J, [[tr]_C q]_D]_K}. \tag{C17}$$

After making the necessary contractions, the total electron-electron repulsion matrix element between SACC states can be rewritten as

$$\langle A, p|\hat{G}|B, q \rangle = S_{pq} G_{AB} + [pq|rs] Q^{BA}_{sr} + [ps|rq] W^{BA}_{sr}$$
$$+ [pt|rs] P^{BA}_{ts,qr} + [qt|rs] P^{BA}_{ps,tr}$$
$$+ [tu|rs] C^{BA}_{psu,trq}. \tag{C18}$$

Notice that the last term involving the third-order TDM needs to be evaluated exclusively when both ions are augmented by an active ionic orbital. Since in this case *all* the orbitals involved are GTOs, the matrix element is most conveniently evaluated within the LUCIA Quantum Chemistry code, which also computes the TDMs.

4. Hamiltonian matrix elements

As in the case of the monoelectronic operators, it is convenient to factor out the contribution of the inactive orbitals in the two-body matrix element. In doing so, it will emerge that the core orbitals contribute to the interelectronic repulsion as an effective one-body potential. The most natural context to separate effective one-body and two-body terms is in the evaluation of the full Hamiltonian, which we presently proceed to carry out.

Here, we assume that the Hamiltonian is diagonal in the ion basis, $\langle A|H|B \rangle = E_A \delta_{AB}$. The matrix element of the total Hamiltonian between SACC states then becomes

$$\langle A, p|\hat{H}|B, q \rangle = \delta_{AB}(S_{pq}E_A + h_{pq})$$
$$+ W^{BA}_{pr}h_{rq} + h_{ps}W^{BA}_{sq} + h_{rs}P^{BA}_{ps,qr}$$
$$+ [pq|rs]Q^{BA}_{sr} + [ps|rq]W^{BA}_{sr}$$
$$+ [pt|rs]P^{BA}_{ts,qr} + [qt|rs]P^{BA}_{ps,tr}$$
$$+ [tu|rs]C^{BA}_{psu,trq}. \tag{C19}$$

All the summations over the ionic orbitals can be restricted to the active orbitals only, provided that the monoelectronic Hamiltonian h_{pq} is replaced by an effective Hamiltonian $\tilde{h}_{pq}$ that incorporates the Coulomb and exchange terms with the

core,

$$\tilde{h}_{pq} = h_{pq} + 2J^{\text{core}}_{pq} - K^{\text{core}}_{pq}, \tag{C20}$$

where

$$J^{\text{core}}_{pq} = \sum_x^{\text{core}} [pq|xx], \quad K^{\text{core}}_{pq} = \sum_x^{\text{core}} [px|xq]. \tag{C21}$$

The total matrix element of the Hamiltonian reads

$$\langle A, p|\hat{H}|B, q \rangle$$
$$= \delta_{AB}(S_{pq}E_A + \tilde{h}_{pq}) + E_{\text{core}} W^{BA}_{pq}$$
$$+ {\sum_r}' (W^{BA}_{pr}\tilde{h}_{rq} + \tilde{h}_{pr}W^{BA}_{rq}) + {\sum_{rs}}' \tilde{h}_{rs} P^{BA}_{ps,qr}$$
$$+ {\sum_{rs}}' ([pq|rs]Q^{BA}_{sr} + [ps|rq]W^{BA}_{sr})$$
$$+ {\sum_{rst}}' ([pt|rs]P^{BA}_{ts,qr} + [qt|rs]P^{BA}_{ps,tr})$$
$$+ {\sum_{turs}}' [tu|rs]C^{BA}_{psu,trq}. \tag{C22}$$

Notice that in this formulation, the matrix elements lend themselves to an efficient implementation. The Hamiltonian matrix elements between active channels can be computed by the LUCIA program. In ASTRA, therefore, (C22) is used only when at least one of the two augmented orbital indexes, p or q, does not correspond to an active internal orbital. Let us assume, therefore, without loss of generality, that q is a virtual orbital (the case in which p is virtual and q is not is obtained by Hermitian conjugation). The expression then greatly simplifies to

$$\langle A, p|\hat{H}|B, q \rangle$$
$$= \delta_{AB}(S_{pq}E_A + \tilde{h}_{pq})$$
$$+ {\sum_r}' W^{BA}_{pr}\tilde{h}_{rq} + {\sum_{rs}}' ([pq|rs]Q^{BA}_{sr} + [ps|rq]W^{BA}_{sr})$$
$$+ {\sum_{rst}}' [qt|rs]P^{BA}_{ps,tr} \quad q \in \text{VO}. \tag{C23}$$

If both p and q are virtual orbitals, the formula simplifies further to

$$\langle A, p|\hat{H}|B, q \rangle = \delta_{AB}(S_{pq}E_A + \tilde{h}_{pq})$$
$$+ {\sum_{rs}}' ([pq|rs]Q^{BA}_{sr} + [ps|rq]W^{BA}_{sr}). \tag{C24}$$

Owing to the compact support character of the spherical numerical functions used to describe the Nth electron at large distances, and to the rapid decrease with r of any of the GTO active orbitals used to build the ions, for all practical purposes, the product of most of the former functions with *any* of the latter is zero. The virtual orbitals with this property are effectively external to the molecular region. If *either* p or q is an external spherical orbital (ESO), then the expression for the Hamiltonian simplifies even further. For example, if q

is an ESO,

$$\langle A, p|\hat{H}|B, q\rangle = \delta_{AB}(S_{pq}E_A + \tilde{h}_{pq})$$
$$+ \sum_{rs}' [pq|rs]Q_{sr}^{BA} \quad q \in \text{ESO.} \quad \text{(C25)}$$

Finally, if q is an AMO and p is an ESO, their supports are effectively disjoint and hence the Hamiltonian matrix element vanishes entirely.

a. Matrix elements between AC blocks

The evaluation with the TDMCC formalism of the Hamiltonian matrix elements between ionic states that are both augmented by active orbitals requires the use of three-body TDMs, which is both computationally and logistically demanding. For this reason, it is convenient to use auxiliary matrix elements that are generated by the same code that prepares the TDMs. LUCIA can generate elements of the form

$$\langle A|a_P \hat{H} a_Q^\dagger|B\rangle, \quad \text{(C26)}$$

in which the ion and electron spins are not coupled. Once a sufficient number of such independent matrix elements are available, it is possible to extract the reduced matrix elements of the associated singlet and triplet operators $[b_p \hat{H} b_q^\dagger]_{T\tau}$, and from these evaluate those between SACC states. Let us define

$$\mathsf{H}_{[p,q]_T}^{AB} \equiv (-1)^{T+1}\langle A\|[b_p\hat{H}b_q^\dagger]_T\|B\rangle$$
$$= (-1)^{T+1}\Pi_{S_A}^{-1}\sum_{\Sigma_A\Sigma_B\pi\theta\tau}(-1)^{\frac{1}{2}-\pi}C_{\frac{1}{2}-\pi,\frac{1}{2}\theta}^{T\tau}$$
$$\times C_{S_B\Sigma_B,T\tau}^{S_A\Sigma_A}\langle A|a_P\hat{H}a_Q^\dagger|B\rangle.$$

The reduced element $\mathsf{H}_{[p,q]_T}^{AB}$ can be determined from the matrix elements between states with a same (suitable) $\Sigma = \Sigma_A = \Sigma_B$, and operators with the same spin projection, $\pi = \theta$,

$$\mathsf{H}_{[p,q]_T}^{AB} = \sum_\theta \frac{\Pi_{S_A}C_{\frac{1}{2}\theta,\frac{1}{2}-\theta}^{T0}}{(-1)^{\frac{1}{2}-\theta}C_{S_B\Sigma,T0}^{S_A\Sigma}}\langle A|a_P\hat{H}a_Q^\dagger|B\rangle.$$

Notice that this is the same relation as the one between $W_{[q,p]_T}^{BA}$ and $\rho_{q_\theta p_\theta}^{BA}$. Once the matrix elements $\mathsf{H}_{[p,q]_T}^{AB}$ are available, the Hamiltonian between SACC states can be shown to be given by

$$\langle A, p|\hat{H}|B, q\rangle = -\eta \sum_T \mathsf{H}_{[p,q]_T}^{AB} \Pi_T \begin{Bmatrix} T & 1/2 & 1/2 \\ S & S_B & S_A \end{Bmatrix}.$$

APPENDIX D: CIS MODEL

ASTRA was tested for consistency by comparison with a CIS model, in which the configuration space comprises a single determinant for the neutral ground state $|\mathbf{K}\rangle$, as well as all the single excitations $|\mathbf{K}_A^P\rangle = a_P^\dagger a_A|\mathbf{K}\rangle$ ($P \neq A$, $\langle P|Q\rangle = \delta_{PQ}$ $\forall P, Q$). This space, of course, coincides with the CC space generated by all the ions $|A\rangle = |\mathbf{K}_A\rangle \equiv a_A|\mathbf{K}\rangle$. All CIS states

are orthonormal,

$$\langle \mathbf{K}_A^P|\mathbf{K}_B^Q\rangle = \delta_{PQ}\delta_{AB}, \quad \langle\mathbf{K}|\mathbf{K}\rangle = 1, \quad \langle\mathbf{K}_A^P|\mathbf{K}\rangle = 0.$$

The matrix elements of arbitrary one-body and two-body operators between such CIS states are readily determined with Slater rules [161]. For monoelectronic operators,

$$\langle \mathbf{K}|\hat{\mathcal{O}}|\mathbf{K}\rangle = \sum_{I\in\mathbf{K}}o_{II} = \langle\hat{\mathcal{O}}\rangle_{\mathbf{K}}, \quad \langle\mathbf{K}|\hat{\mathcal{O}}|\mathbf{K}_B^Q\rangle = o_{AQ},$$

$$\langle\mathbf{K}_A^P|\hat{\mathcal{O}}|\mathbf{K}_B^Q\rangle = \delta_{AB}[\delta_{PQ}(\langle\hat{\mathcal{O}}\rangle_{\mathbf{K}} - o_{AA}) + o_{PQ}].$$

Let us introduce the notation $\langle IJ\|KL\rangle = \langle IJ|KL\rangle - \langle IJ|LK\rangle = [IK|JL] - [IL|JK]$. For the matrix elements between CIS states of the interelectronic repulsion then are

$$\langle\mathbf{K}|\hat{G}|\mathbf{K}\rangle = \frac{1}{2}\sum_{I,J\in\mathbf{K}}\langle IJ\|IJ\rangle,$$

$$\langle\mathbf{K}|\hat{G}|\mathbf{K}_B^Q\rangle = \sum_{I\in\mathbf{K}}\langle BI\|QI\rangle,$$

$$\langle\mathbf{K}_A^P|\hat{G}|\mathbf{K}_A^P\rangle = \frac{1}{2}\sum_{I,J\in\mathbf{K}_A^P}\langle IJ\|IJ\rangle,$$

$$\langle\mathbf{K}_A^P|\hat{G}|\mathbf{K}_A^{Q\neq P}\rangle = \sum_{I\in\mathbf{K}}\langle AI\|QI\rangle + \langle AP\|QP\rangle,$$

$$\langle\mathbf{K}_A^P|\hat{G}|\mathbf{K}_{B\neq A}^P\rangle = \sum_{I\in\mathbf{K}}\langle BI\|AI\rangle + \langle BP\|AP\rangle,$$

$$\langle\mathbf{K}_A^P|\hat{G}|\mathbf{K}_{B\neq A}^{Q\neq P}\rangle = \langle BP\|AQ\rangle.$$

If the initial state is a closed-shell Hartree-Fock state, then the expressions can be simplified into sums over the doubly occupied orbitals. These results are sufficient to implement an *ad hoc* code to test the accuracy of the results obtained with the more sophisticated ASTRA code, in the CIS limit.

Thanks to its simplicity, the CIS approximation offers also a way of testing the consistency of the phase conventions in the interface between LUCIA and ASTRA. Indeed, within the CIS approximation, the TDM have the following elementary expressions

$$\rho_{QP}^{BA} = \langle\mathbf{K}|a_A^\dagger a_P^\dagger a_Q a_B|\mathbf{K}\rangle$$
$$= (1 - \delta_{AP})(1 - \delta_{BQ})(\delta_{AB}\delta_{PQ} - \delta_{AQ}\delta_{PB}),$$

$$\pi_{RS,PQ}^{BA} = \langle\mathbf{K}|a_A^\dagger a_P^\dagger a_Q^\dagger a_S a_R a_B|\mathbf{K}\rangle$$
$$= (1 - \delta_{AP})(1 - \delta_{AQ})(1 - \delta_{PQ})$$
$$\times (1 - \delta_{BR})(1 - \delta_{BS})(1 - \delta_{RS})$$
$$\times (\delta_{AB}\delta_{PR}\delta_{QS} - \delta_{AB}\delta_{QR}\delta_{PS} + \delta_{QB}\delta_{AR}\delta_{PS}$$
$$- \delta_{PB}\delta_{AR}\delta_{QS} + \delta_{PB}\delta_{QR}\delta_{AS} - \delta_{QB}\delta_{PR}\delta_{AS}).$$

[1] P. M. Kraus and H. J. Wörner, Perspectives of attosecond spectroscopy for the understanding of fundamental electron correlations, Angew. Chem. Int. Ed. **57**, 5228 (2018).

[2] M. Hentschel, R. Kienberger, C. Spielmann, G. A. Reider, N. Milosevic, T. Brabec, P. B Corkum, U. Heinzmann, M. Drescher, and F. Krausz, Attosecond metrology, Nature (London) **414**, 509 (2001).

[3] P. M. Paul, E. S. Toma, P. Breger, G. Mullot, F. Augé, P. Balcou, H. G. Muller, and P. Agostini, Observation of a train of attosecond pulses from high harmonic generation, Science **292**, 1689 (2001).

[4] F. Krausz and M. Yu. Ivanov, Attosecond physics, Rev. Mod. Phys. **81**, 163 (2009).

[5] M. Nisoli, P. Decleva, F. Calegari, A. Palacios, and F. Martín, Attosecond electron dynamics in molecules, Chem. Rev. **117**, 10760 (2017).

[6] A. Palacios and F. Martín, The quantum chemistry of attosecond molecular science, WIREs Comput. Mol. Sci. **10**, e1430 (2020).

[7] R. Borrego-Varillas, M. Lucchini, and M. Nisoli, Attosecond spectroscopy for the investigation of ultrafast dynamics in atomic, molecular and solid-state physics, Rep. Prog. Phys. **85**, 066401 (2022).

[8] S. R. Leone and D. M. Neumark, Probing matter with nonlinear spectroscopy, Science **379**, 536 (2023).

[9] J. Feldhaus, M. Krikunova, M. Meyer, Th. Möller, R. Moshammer, A. Rudenko, Th. Tschentscher, and J. Ullrich, AMO science at the FLASH and European XFEL free-electron laser facilities, J. Phys. B: At. Mol. Opt. Phys. **46**, 164002 (2013).

[10] N. Hartmann, G. Hartmann, R. Heider, M. S. Wagner, M. Ilchen, J. Buck, A. O. Lindahl, C. Benko, J. Grünert, J. Krzywinski *et al.*, Attosecond time-energy structure of x-ray free-electron laser pulses, Nat. Photonics **12**, 215 (2018).

[11] T. Popmintchev, M.-C. Chen, P. Arpin, M. M. Murnane, and H. C. Kapteyn, The attosecond nonlinear optics of bright coherent x-ray generation, Nat. Photonics **4**, 822 (2010).

[12] G. Sansone, L. Poletto, and M. Nisoli, High-energy attosecond light sources, Nat. Photonics **5**, 655 (2011).

[13] M. Chini, K. Zhao, and Z. Chang, The generation, characterization and applications of broadband isolated attosecond pulses, Nat. Photonics **8**, 178 (2014).

[14] F. Calegari, G. Sansone, S. Stagira, C. Vozzi, and M. Nisoli, Advances in attosecond science, J. Phys. B: At. Mol. Opt. Phys. **49**, 062001 (2016).

[15] S. Kühn, M. Dumergue, S. Kahaly, S. Mondal, M. Füle, T. Csizmadia, B. Farkas, B. Major, Z. Várallyay, E. Cormier *et al.*, The ELI-ALPS facility: The next generation of attosecond sources, J. Phys. B: At. Mol. Opt. Phys. **50**, 132002 (2017).

[16] K. Midorikawa, Progress on table-top isolated attosecond light sources, Nat. Photonics **16**, 267 (2022).

[17] N. Saito, H. Sannohe, N. Ishii, T. Kanai, N. Kosugi, Y. Wu, A. Chew, S. Han, Z. Chang, and J. Itatani, Real-time observation of electronic, vibrational, and rotational dynamics in nitric oxide with attosecond soft x-ray pulses at 400 eV, Optica **6**, 1542 (2019).

[18] J. Duris, S. Li, T. Driver, E. G. Champenois, J. P. MacArthur, A. A. Lutman, Z. Zhang, P. Rosenberger, J. W. Aldrich, R. Coffee *et al.*, Tunable isolated attosecond x-ray pulses with gigawatt peak power from a free-electron laser, Nat. Photonics **14**, 30 (2020).

[19] C. Callegari, A. N. Grum-Grzhimailo, K. L. Ishikawa, K. C. Prince, G. Sansone, and K. Ueda, Atomic, molecular and optical physics applications of longitudinally coherent and narrow bandwidth free-electron lasers, Phys. Rep. **904**, 1 (2021).

[20] N. S. Mirian, M. Di Fraia, S. Spampinati, F. Sottocorona, E. Allaria, L. Badano, M. B. Danailov, A. Demidovich, G. De Ninno, S. Di Mitri *et al.*, Generation and measurement of intense few-femtosecond superradiant extreme-ultraviolet free-electron laser pulses, Nat. Photonics **15**, 523 (2021).

[21] F. Calegari, D. Ayuso, A. Trabattoni, L. Belshaw, S. De Camillis, S. Anumula, F. Frassetto, L. Poletto, A. Palacios, P. Decleva *et al.*, Ultrafast electron dynamics in phenylalanine initiated by attosecond pulses, Science **346**, 336 (2014).

[22] K. Ueda, E. Sokell, S. Schippers, F. Aumayr, H. Sadeghpour, J. Burgdörfer, C. Lemell, X.-M. Tong, T. Pfeifer, F. Calegari *et al.*, Roadmap on photonic, electronic and atomic collision physics: I. Light-matter interaction, J. Phys. B: At. Mol. Opt. Phys. **52**, 171001 (2019).

[23] E. P. Månsson, S. Latini, F. Covito, V. Wanie, M. Galli, E. Perfetto, G. Stefanucci, H. Hübener, U. De Giovannini, M. C. Castrovilli *et al.*, Real-time observation of a correlation-driven sub 3 fs charge migration in ionised adenine, Commun. Chem. **4**, 73 (2021).

[24] G. S. J. Armstrong, M. A. Khokhlova, M. Labeye, A. S. Maxwell, E. Pisanty, and M. Ruberti, Dialogue on analytical and *ab initio* methods in attoscience, Eur. Phys. J. D **75**, 209 (2021).

[25] J. Olsen, LUCIA, a general CI code written by J. Olsen, University of Aarhus with contributions from H. Larsen and M. Fülscher (2000).

[26] P. G. Burke and H. M Schey, Elastic scattering of low-energy electrons by atomic hydrogen, Phys. Rev. **126**, 147 (1962).

[27] B. Rudek, S.-k. Son, L. Foucar, S. W. Epp, B. Erk, R. Hartmann, M. Adolph, R. Andritschke, A. Aquila, N. Berrah *et al.*, Ultra-efficient ionization of heavy atoms by intense x-ray free-electron laser pulses, Nat. Photonics **6**, 858 (2012).

[28] Y. Hao, L. Inhester, S.-K. Son, and R. Santra, Theoretical evidence for the sensitivity of charge-rearrangement-enhanced x-ray ionization to molecular size, Phys. Rev. A **100**, 013402 (2019).

[29] A. Rudenko, L. Inhester, K. Hanasaki, X. Li, S. J. Robatjazi, B. Erk, R. Boll, K. Toyota, Y. Hao, O. Vendrell *et al.*, Femtosecond response of polyatomic molecules to ultra-intense hard x-rays, Nature (London) **546**, 129 (2017).

[30] A. P. P. Natalense and R. R. Lucchese, Cross section and asymmetry parameter calculation for sulfur 1s photoionization of SF_6, J. Chem. Phys. **111**, 5344 (1999).

[31] X. Gong, W. Jiang, J. Tong, J. Qiang, P. Lu, H. Ni, R. Lucchese, K. Ueda, and J. Wu, Asymmetric attosecond photoionization in molecular shape resonance, Phys. Rev. X **12**, 011002 (2022).

[32] M. Venuti, M. Stener, and P. Decleva, Valence photoionization of C_6H_6 by the B-spline one-centre expansion density functional method, Chem. Phys. **234**, 95 (1998).

[33] M. Stener and P. Decleva, Photoionization of first and second row hydrides by the B-spline one-centre expansion density functional method, J. Electron Spectrosc. Relat. Phenom. **94**, 195 (1998).

[34] D. Toffoli, M. Stener, G. Fronzoni, and P. Decleva, Convergence of the multicenter B-spline DFT

approach for the continuum, Chem. Phys. **276**, 25 (2002).

[35] P. Decleva, M. Stener, and D. Toffoli, Continuum electronic states: The Tiresia code, Molecules **27**, 2026 (2022).

[36] S. Petretti, Y. V. Vanne, A. Saenz, A. Castro, and P. Decleva, Alignment-dependent ionization of N_2O_2, and CO, Phys. Rev. Lett. **104**, 223001 (2010).

[37] S. E. Canton, E. Plésiat, J. D. Bozek, B. S. Rude, P. Decleva, and F. Martín, Direct observation of Young's double-slit interferences in vibrationally resolved photoionization of diatomic molecules, Proc. Natl. Acad. Sci. USA **108**, 7302 (2011).

[38] E. Plésiat, L. Argenti, E. Kukk, C. Miron, K. Ueda, P. Decleva, and F. Martín, Intramolecular electron diffraction in vibrationally resolved K-shell photoionization of methane, Phys. Rev. A **85**, 023409 (2012).

[39] L. Argenti, T. D. Thomas, E. Plésiat, X.-J. Liu, C. Miron, T. Lischke, G. Prümper, K. Sakai, T. Ouchi, R. Püttner *et al.*, Double-slit experiment with a polyatomic molecule: Vibrationally resolved C 1s photoelectron spectra of acetylene, New J. Phys. **14**, 033012 (2012).

[40] M. Lara-Astiaso, M. Galli, A. Trabattoni, A. Palacios, D. Ayuso, F. Frassetto, L. Poletto, S. De Camillis, J. Greenwood, P. Decleva *et al.*, Attosecond pump-probe spectroscopy of charge dynamics in tryptophan, J. Phys. Chem. Lett. **9**, 4570 (2018).

[41] O. Neufeld, D. Ayuso, P. Decleva, M. Y. Ivanov, O. Smirnova, and O. Cohen, Ultrasensitive chiral spectroscopy by dynamical symmetry breaking in high harmonic generation, Phys. Rev. X **9**, 031002 (2019).

[42] U. De Giovannini, D. Varsano, M. A. L. Marques, H. Appel, E. K. U. Gross, and A. Rubio, *Ab initio* angle- and energy-resolved photoelectron spectroscopy with time-dependent density-functional theory, Phys. Rev. A **85**, 062515 (2012).

[43] U. De Giovannini, G. Brunetto, A. Castro, J. Walkenhorst, and A. Rubio, Simulating pump-probe photoelectron and absorption spectroscopy on the attosecond timescale with time-dependent density functional theory, ChemPhysChem **14**, 1363 (2013).

[44] M. Awasthi, Y. V. Vanne, A. Saenz, A. Castro, and P. Decleva, Single-active-electron approximation for describing molecules in ultrashort laser pulses and its application to molecular hydrogen, Phys. Rev. A **77**, 063403 (2008).

[45] E. Kukk, D. Ayuso, T. D. Thomas, P. Decleva, M. Patanen, L. Argenti, E. Plésiat, A. Palacios, K. Kooser, O. Travnikova, S. Mondal, M. Kimura, K. Sakai, C. Miron, F. Martín, and K. Ueda, Effects of molecular potential and geometry on atomic core-level photoemission over an extended energy range: The case study of the CO molecule, Phys. Rev. A **88**, 033412 (2013).

[46] D. Ayuso, A. Palacios, P. Decleva, and F. Martín, Dissociative and non-dissociative photoionization of molecular fluorine from inner and valence shells, J. Electron Spectrosc. Relat. Phenom. **195**, 320 (2014).

[47] A. U. J. Lode, C. Lévêque, L. B. Madsen, A. I. Streltsov, and O. E. Alon, Colloquium: Multiconfigurational time-dependent Hartree approaches for indistinguishable particles, Rev. Mod. Phys. **92**, 011001 (2020).

[48] D. J. Haxton, K. V. Lawler, and C. W. McCurdy, Multiconfiguration time-dependent Hartree-Fock treatment of electronic

and nuclear dynamics in diatomic molecules, Phys. Rev. A **83**, 063416 (2011).

[49] R. Sawada, T. Sato, and K. L. Ishikawa, Implementation of the multiconfiguration time-dependent Hatree-Fock method for general molecules on a multiresolution Cartesian grid, Phys. Rev. A **93**, 023434 (2016).

[50] C.-t. Liao, X. Li, D. J. Haxton, T. N. Rescigno, R. R. Lucchese, C. W. McCurdy, and A. Sandhu, Probing autoionizing states of molecular oxygen with XUV transient absorption: Electronic-symmetry-dependent line shapes and laser-induced modifications, Phys. Rev. A **95**, 043427 (2017).

[51] H. Miyagi and L. B. Madsen, Time-dependent restricted-active-space self-consistent-field theory for laser-driven many-electron dynamics, Phys. Rev. A **87**, 062511 (2013).

[52] J. J. Omiste and L. B. Madsen, Photoionization of aligned excited states in neon by attosecond laser pulses, J. Phys. B: At. Mol. Opt. Phys. **54**, 054001 (2021).

[53] S. Bauch, L. K. Sørensen, and L. B. Madsen, Time-dependent generalized-active-space configuration-interaction approach to photoionization dynamics of atoms and molecules, Phys. Rev. A **90**, 062508 (2014).

[54] S. Chattopadhyay, S. Bauch, and L. B. Madsen, Electron-correlation effects in enhanced ionization of molecules: A time-dependent generalized-active-space configuration-interaction study, Phys. Rev. A **92**, 063423 (2015).

[55] T. Sato and K. L. Ishikawa, Time-dependent complete-active-space self-consistent-field method for multielectron dynamics in intense laser fields, Phys. Rev. A **88**, 023402 (2013).

[56] Y. Li, T. Sato, and K. L. Ishikawa, Implementation of a time-dependent multiconfiguration self-consistent-field method for coupled electron-nuclear dynamics in diatomic molecules driven by intense laser pulses, Phys. Rev. A **104**, 043104 (2021).

[57] T. Sato, H. Pathak, Y. Orimo, and K. L. Ishikawa, Communication: Time-dependent optimized coupled-cluster method for multielectron dynamics, J. Chem. Phys. **148**, 051101 (2018).

[58] D. Toffoli and P. Decleva, A multichannel least-squares B-spline approach to molecular photoionization: Theory, implementation, and applications within the configuration-interaction singles approximation, J. Chem. Theory Comput. **12**, 4996 (2016).

[59] P. Hoerner, W. Li, and H. B. Schlegel, Angular dependence of strong field ionization of 2-phenylethyl- N, N -dimethylamine (PENNA) using time-dependent configuration interaction with an absorbing potential, J. Phys. Chem. A **124**, 4777 (2020).

[60] S. Carlström, M. Spanner, and S. Patchkovskii, General time-dependent configuration-interaction singles. I. Molecular case, Phys. Rev. A **106**, 043104 (2022).

[61] H. Bachau, E. Cormier, P. Decleva, J. E. Hansen, and F. Martín, Applications of B -splines in atomic and molecular physics, Rep. Prog. Phys. **64**, 1815 (2001).

[62] F. L. Yip, C. W. McCurdy, and T. N. Rescigno, Hybrid Gaussian-discrete-variable representation approach to molecular continuum processes: Application to photoionization of diatomic Li_2^+, Phys. Rev. A **78**, 023405 (2008).

[63] L. Greenman, R. R. Lucchese, and C. W. McCurdy, Variational treatment of electron-polyatomic-molecule scattering calculations using adaptive overset grids, Phys. Rev. A **96**, 052706 (2017).

[64] P. G. Burke, I. Mackey, and I. Shimamura, R-matrix theory of electron-molecule scattering, J. Phys. B: At. Mol. Phys. **10**, 2497 (1977).

[65] J. D. Gorfinkiel and J. Tennyson, Electron impact ionization of small molecules at intermediate energies: The molecular R-matrix with pseudostates method, J. Phys. B: At. Mol. Opt. Phys. **38**, 1607 (2005).

[66] A. G. Harvey, D. S. Brambila, F. Morales, and O. Smirnova, An R-matrix approach to electron-photon-molecule collisions: Photoelectron angular distributions from aligned molecules, J. Phys. B: At. Mol. Opt. Phys. **47**, 215005 (2014).

[67] Z. Mašín, J. Benda, J. D. Gorfinkiel, A. G. Harvey, and J. Tennyson, UKRmol+: A suite for modelling electronic processes in molecules interacting with electrons, positrons and photons using the R-matrix method, Comput. Phys. Commun. **249**, 107092 (2020).

[68] A. C. Brown, G. S. J. Armstrong, J. Benda, D. D. A. Clarke, J. Wragg, K. R. Hamilton, Z. Mašín, J. D. Gorfinkiel, and H. W. van der Hart, RMT: R-matrix with time-dependence. Solving the semi-relativistic, time-dependent Schrödinger equation for general, multielectron atoms and molecules in intense, ultrashort, arbitrarily polarized laser pulses, Comput. Phys. Commun. **250**, 107062 (2020).

[69] I. Cacelli, The calculation of photoionisation cross sections of simple polyatomic molecules by L2 methods, Phys. Rep. **205**, 283 (1991).

[70] I. Cacelli, R. Moccia, and R. Montuoro, Calculation of the differential photoionization cross-section of formaldehyde, Chem. Phys. Lett. **347**, 261 (2001).

[71] B. I. Schneider and T. N. Rescigno, Complex Kohn variational method: Application to low-energy electron-molecule collisions, Phys. Rev. A **37**, 3749 (1988).

[72] C. W. McCurdy and T. N. Rescigno, Collisions of electrons with polyatomic molecules: Electron-methane scattering by the complex Kohn variational method, Phys. Rev. A **39**, 4487 (1989).

[73] B. I. Schneider, T. N. Rescigno, and C. W. McCurdy, Resonant vibrational excitation of H$_2$CO by low-energy electron impact, Phys. Rev. A **42**, 3132 (1990).

[74] T. N. Rescigno, B. H. Lengsfield, and A. E. Orel, Interchannel coupling and ground state correlation effects in the photoionization of CO, J. Chem. Phys. **99**, 5097 (1993).

[75] J. B. Williams, C. S. Trevisan, M. S. Schöffler, T. Jahnke, I. Bocharova, H. Kim, B. Ulrich, R. Wallauer, F. Sturm, T. N. Rescigno *et al.*, Probing the dynamics of dissociation of methane following core ionization using three-dimensional molecular-frame photoelectron angular distributions, J. Phys. B: At. Mol. Opt. Phys. **45**, 194003 (2012).

[76] N. Douguet, B. I. Schneider, and L. Argenti, Application of the complex Kohn variational method to attosecond spectroscopy, Phys. Rev. A **98**, 023403 (2018).

[77] V. P. Majety and A. Scrinzi, Static field ionization rates for multi-electron atoms and small molecules, J. Phys. B: At. Mol. Opt. Phys. **48**, 245603 (2015).

[78] A. Scrinzi, tRecX - An environment for solving time-dependent Schrödinger-like problems, Comput. Phys. Commun. **270**, 108146 (2022).

[79] C. Marante, M. Klinker, T. Kjellsson, E. Lindroth, J. González-Vázquez, L. Argenti, and F. Martín, Photoionization using the XCHEM approach: Total and partial cross sections of Ne and resonance parameters above the $2s^2 2p^5$ threshold, Phys. Rev. A **96**, 022507 (2017).

[80] C. Marante, M. Klinker, I. Corral, J. González-Vázquez, L. Argenti, and F. Martín, Hybrid-basis close-coupling interface to quantum chemistry packages for the treatment of ionization problems, J. Chem. Theory Comput. **13**, 499 (2017).

[81] M. Klinker, C. Marante, L. Argenti, J. González-Vázquez, and F. Martín, Electron correlation in the ionization continuum of molecules: Photoionization of N$_2$ in the vicinity of the Hopfield series of autoionizing states, J. Phys. Chem. Lett. **9**, 756 (2018).

[82] M. Klinker, C. Marante, L. Argenti, J. González-Vázquez, and F. Martín, Partial cross sections and interfering resonances in photoionization of molecular nitrogen, Phys. Rev. A **98**, 033413 (2018).

[83] S. Marggi Poullain, M. Klinker, J. González-Vázquez, and F. Martín, Resonant photoionization of O$_2$ up to the fourth ionization threshold, Phys. Chem. Chem. Phys. **21**, 16497 (2019).

[84] V. J. Borràs, J. González-Vázquez, L. Argenti, and F. Martín, Molecular-frame photoelectron angular distributions of CO in the vicinity of Feshbach resonances: An XCHEM approach, J. Chem. Theory Comput. **17**, 6330 (2021).

[85] V. J. Borràs, J. González-Vázquez, L. Argenti, and F. Martín, Attosecond photoionization delays in the vicinity of molecular Feshbach resonances, Sci. Adv. **9**, eade3855 (2023).

[86] M. Ruberti, P. Decleva, and V. Averbukh, Full *ab initio* many-electron simulation of attosecond molecular pump-probe spectroscopy, J. Chem. Theory Comput. **14**, 4991 (2018).

[87] M. Ruberti, Restricted correlation space B-spline ADC approach to molecular ionization: Theory and applications to total photoionization cross-sections, J. Chem. Theory Comput. **15**, 3635 (2019).

[88] G. Bandarage and R. R. Lucchese, Multiconfiguration multichannel Schwinger study of the C(1s) photoionization of CO including shake-up satellites, Phys. Rev. A **47**, 1989 (1993).

[89] R. E. Stratmann and R. R. Lucchese, A graphical unitary group approach to study multiplet specific multichannel electron correlation effects in the photoionization of O$_2$, J. Chem. Phys. **102**, 8493 (1995).

[90] R. R. Lucchese, J. Söderström, T. Tanaka, M. Hoshino, M. Kitajima, H. Tanaka, A. De Fanis, J. E. Rubensson, and K. Ueda, Vibrationally resolved partial cross sections and asymmetry parameters for nitrogen K-shell photoionization of the N$_2$O molecule, Phys. Rev. A **76**, 012506 (2007).

[91] M. Hamermesh, *Group Theory and its Application to Physical Problems*, 2nd ed. (Dover, Reading, MA, 1989).

[92] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic-Structure Theory* (Wiley, Hoboken, NJ, 2000) Chap. 1, pp. 1–33.

[93] R. McWeeny and B. T. Sutcliffe, *Methods of Molecular Quantum Mechanics*, 2nd ed. (Academic Press, London, 1969).

[94] T. N. Rescigno and C. W. McCurdy, Numerical grid methods for quantum-mechanical scattering problems, Phys. Rev. A **62**, 032706 (2000).

[95] D. A. Varshalovich, A. N. Moskalev, and V. K. Khersonskii, *Quantum Theory of Angular Momentum* (World Scientific, Singapore, 1988).

[96] J. Muga, J. Palao, B. Navarro, and I. Egusquiza, Complex absorbing potentials, Phys. Rep. **395**, 357 (2004).

[97] A. J. F. Siegert, On the derivation of the dispersion formula for nuclear reactions, Phys. Rev. **56**, 750 (1939).

[98] C. W. McCurdy and T. N. Rescigno, Basis-set calculation of Siegert eigenvalues: Partial resonance widths, Phys. Rev. A **20**, 2346 (1979).

[99] U. V. Riss and H. D. Meyer, Calculation of resonance energies and widths using the complex absorbing potential method, J. Phys. B: At. Mol. Opt. Phys. **26**, 4503 (1993).

[100] R. G. Newton, *Scattering Theory of Waves and Particles*, Theoretical and Mathematical Physics (Springer, Berlin, 2013).

[101] J. R. Taylor, *Scattering Theory: The Quantum Theory of Non-relativistic Collisions*, Dover Books on Engineering (Dover Publications, 2012).

[102] K. Aidas, C. Angeli, K. L. Bak, V. Bakken, R. Bast, L. Boman, O. Christiansen, R. Cimiraglia, S. Coriani, P. Dahle *et al.*, The Dalton quantum chemistry program system, WIREs Comput. Mol. Sci. **4**, 269 (2014).

[103] J. Olsen, B. O. Roos, P. Jørgensen, and H. J. A. Jensen, Determinant based configuration interaction algorithms for complete and restricted configuration interaction spaces, J. Chem. Phys. **89**, 2185 (1988).

[104] P. J. Knowles and N. C. Handy, A new determinant-based full configuration interaction method, Chem. Phys. Lett. **111**, 315 (1984).

[105] J. Olsen, The initial implementation and applications of a general active space coupled cluster method, J. Chem. Phys. **113**, 7140 (2000).

[106] S. Kähler and J. Olsen, Non-orthogonal internally contracted multi-configurational perturbation theory (NICPT): Dynamic electron correlation for large, compact active spaces, J. Chem. Phys. **147**, 174106 (2017).

[107] M. Brosolo, P. Decleva, and A. Lisini, Continuum wavefunctions calculations with least-squares schemes in a B-splines basis, Comput. Phys. Commun. **71**, 207 (1992).

[108] C. Marante, L. Argenti, and F. Martín, Hybrid Gaussian-B-spline basis for the electronic continuum: Photoionization of atomic hydrogen, Phys. Rev. A **90**, 012506 (2014).

[109] H. H. H. Homeier and E. O. Steinborn, Some properties of the coupling coefficients of real spherical harmonics and their relation to Gaunt coefficients, J. Mol. Struct (THEOCHEM) **368**, 31 (1996).

[110] T. Carette, J. M. Dahlström, L. Argenti, and E. Lindroth, Multiconfigurational Hartree-Fock close-coupling ansatz: Application to the argon photoionization cross section and delays, Phys. Rev. A **87**, 023420 (2013).

[111] L. Argenti and R. Moccia, Autoionizing states of atomic boron, Phys. Rev. A **93**, 042503 (2016).

[112] C. Froese Fischer, G. Tachiev, G. Gaigalas, and M. R. Godefroid, An MCHF atomic-structure package for large-scale calculations, Comput. Phys. Commun. **176**, 559 (2007).

[113] A. Kramida, Y. Ralchenko, and J. Reader, NIST Atomic Spectra Database (version 5.10), https://www.nist.gov/pml/atomic-spectra-database (2022).

[114] C. Froese Fischer, S. Verdebout, M. Godefroid, P. Rynkun, P. Jönsson, and G. Gaigalas, Doublet-quartet energy separation in boron: A partitioned-correlation-function-interaction method, Phys. Rev. A **88**, 062506 (2013).

[115] E. R. Davidson, S. A. Hagstrom, S. J. Chakravorty, V. M. Umar, and C. F. Fischer, Ground-state correlation energies for two- to ten-electron atomic ions, Phys. Rev. A **44**, 7071 (1991).

[116] B. Simon, The definition of molecular resonance curves by the method of exterior complex scaling, Phys. Lett. A **71**, 211 (1979).

[117] E. Lindroth, Photodetachment of H^- and Li^-, Phys. Rev. A **52**, 2737 (1995).

[118] C. W. McCurdy and F. Martín, Implementation of exterior complex scaling in B-splines to solve atomic and molecular collision problems, J. Phys. B: At. Mol. Opt. Phys. **37**, 917 (2004).

[119] N. G. Puskar, Y.-C. Lin, J. D. Gaynor, M. C. Schuchter, S. Chattopadhyay, C. Marante, A. P. Fidler, C. L. Keenan, L. Argenti, D. M. Neumark, and S. R. Leone, Measuring autoionization decay lifetimes of optically forbidden inner valence excited states in neon atoms with attosecond noncollinear four-wave-mixing spectroscopy, Phys. Rev. A **107**, 033117 (2023).

[120] K. P. Huber and G. Herzberg, *Molecular Spectra and Molecular Structure* (Springer US, Boston, MA, 1979).

[121] P. Gürtler, V. Saile, and E. E. Koch, High resolution absorption spectrum of nitrogen in the vacuum ultraviolet, Chem. Phys. Lett. **48**, 245 (1977).

[122] W. B. Peatman, B. Gotchev, P. Gürtler, E. E. Koch, and V. Saile, Transition probabilities at threshold for the photoionization of molecular nitrogen, J. Chem. Phys. **69**, 2089 (1978).

[123] P. M. Dehmer, P. J. Miller, and W. A. Chupka, Photoionization of N_2 X $^1\Sigma_g^+$, $v'' = 0$ and 1 near threshold. Preionization of the Worley-Jenkins Rydberg series, J. Chem. Phys. **80**, 1030 (1984).

[124] K. P. Huber, G. Stark, and K. Ito, Rotational structure in the Hopfield series of N_2, J. Chem. Phys. **98**, 4471 (1993).

[125] P. Süle, O. V. Gritsenko, Á. Nagy, and E. J. Baerends, Correlation energy density from *ab initio* first- and second-order density matrices: A benchmark for approximate functionals, J. Chem. Phys. **103**, 10085 (1995).

[126] H. Feshbach, A unified theory of nuclear reactions. II, Ann. Phys. **19**, 287 (1962).

[127] P. Saxe, B. H. Lengsfield, R. Martin and M. Page, MESA (Molecular Electronic Structure Applications), a software package developed at Los Alamos National Lab. (1990).

[128] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen, *LAPACK Users' Guide*, 3rd ed. (Society for Industrial and Applied Mathematics, Philadelphia, PA, 1999).

[129] G. Cooper, J. E Anderson, and C. E. Brion, Absolute photoabsorption and photoionization of formaldehyde in the VUV and soft x-ray regions (3-200 eV), Chem. Phys. **209**, 61 (1996).

[130] H. K. Tanaka, F. V. Prudente, A. Medina, R. R. T. Marinho, M. G. P. Homem, L. E. Machado, and M. M. Fujimoto, Photoabsorption and photoionization cross sections for formaldehyde in the vacuum-ultraviolet energy range, J. Chem. Phys. **146**, 094310 (2017).

[131] K. Wang, Y. Wang, H. Qiao, J. Liu, H. Zhang, and G. Li, Outer valence photoionization and autoionization of formaldehyde, J. Phys. Chem. A **126**, 5559 (2022).

[132] L. V. Gurvich, I. V. Veyts, and C. B. Alcock, *Thermodynamic Properties of Individual Substances* (CRC Press, Boca Raton, FL, 1989).

[133] G. Schaftenaar and J. H. Noordik, Molden: A pre- and post-processing program for molecular and electronic structures, J. Comput.-Aided Mol. Des. **14**, 123 (2000).

[134] G. Schaftenaar, E. Vlieg, and G. Vriend, Molden 2.0: Quantum chemistry meets proteins, J. Comput.-Aided Mol. Des. **31**, 789 (2017).

[135] G. Wu, P. Hockett, and A. Stolow, Time-resolved photoelectron spectroscopy: From wavepackets to observables, Phys. Chem. Chem. Phys. **13**, 18447 (2011).

[136] D. Dowek and P. Decleva, Trends in angle-resolved molecular photoelectron spectroscopy, Phys. Chem. Chem. Phys. **24**, 24614 (2022).

[137] P. A. Malmqvist, A. Rendell, and B. O. Roos, The restricted active space self-consistent-field method, implemented with a split graph unitary group approach, J. Phys. Chem. **94**, 5477 (1990).

[138] H. Lischka, D. Nachtigallová, A. J. A. Aquino, P. G. Szalay, F. Plasser, F. B. C. Machado, and M. Barbatti, Multireference approaches for excited states of molecules, Chem. Rev. **118**, 7293 (2018).

[139] C. Zhou, L. Gagliardi, and D. G. Truhlar, Multiconfiguration pair-density functional theory for iron porphyrin with CAS, RAS, and DMRG active spaces, J. Phys. Chem. A **123**, 3389 (2019).

[140] D. Casanova, Restricted active space configuration interaction methods for strong correlation: Recent developments, WIREs Comput. Mol. Sci. **12**, e1561 (2022).

[141] N. Harkema, C. Cariker, E. Lindroth, L. Argenti, and A. Sandhu, Autoionizing polaritons in attosecond atomic ionization, Phys. Rev. Lett. **127**, 023202 (2021).

[142] V. Leshchenko, S. J. Hageman, C. Cariker, G. Smith, A. Camper, B. K. Talbert, P. Agostini, L. Argenti, and L. F. DiMauro, Kramers-Kronig relation in attosecond transient absorption spectroscopy, Optica **10**, 142 (2023).

[143] Y. Kobayashi, K. F Chang, T. Zeng, D. M. Neumark, and S. R. Leone, Direct mapping of curve-crossing dynamics in IBr by attosecond transient absorption spectroscopy, Science **365**, 79 (2019).

[144] P. M. Kraus, M. Zürch, S. K. Cushing, D. M. Neumark, and S. R. Leone, The ultrafast x-ray spectroscopic revolution in chemical dynamics, Nat. Rev. Chem. **2**, 82 (2018).

[145] K. S. Zinchenko, F. Ardana-Lamas, I. Seidu, S. P. Neville, J. van der Veen, V. Utrio Lanfaloni, M. S. Schuurman, and H. J. Wörner, Sub-7-femtosecond conical-intersection dynamics probed at the carbon K-edge, Science **371**, 489 (2021).

[146] M. Rebholz, T. Ding, V. Despré, L. Aufleger, M. Hartmann, K. Meyer, V. Stooß, A. Magunia, D. Wachs, P. Birk, Y. Mi, G. D. Borisova, C. C. Castanheira, P. Rupprecht, G. Schmid, K. Schnorr, C. D. Schröter, R. Moshammer, Z.-H. Loh, A. R. Attar *et al.*, All-XUV pump-probe transient absorption spectroscopy of the structural molecular dynamics of di-iodomethane, Phys. Rev. X **11**, 031001 (2021).

[147] W. Cao, E. R. Warrick, A. Fidler, S. R. Leone, and D. M. Neumark, Excited-state vibronic wave-packet dynamics in H_2 probed by XUV transient four-wave mixing, Phys. Rev. A **97**, 023401 (2018).

[148] E. R. Warrick, A. P. Fidler, W. Cao, E. Bloch, D. M. Neumark, and S. R. Leone, Multiple pulse coherent dynamics and wave packet control of the $N_2 a''^1\Sigma_g^+$ dark state by attosecond four-wave mixing, Faraday Discuss. **212**, 157 (2018).

[149] A. P Fidler, Y.-C. Lin, J. D Gaynor, C. W. McCurdy, S. R. Leone, R. R. Lucchese, and D. M. Neumark, State-selective probing of CO_2 autoionizing inner valence Rydberg states with attosecond extreme ultraviolet four-wave-mixing spectroscopy, Phys. Rev. A **106**, 063525 (2022).

[150] M. S. Schuurman and V. Blanchet, Time-resolved photoelectron spectroscopy: The continuing evolution of a mature technique, Phys. Chem. Chem. Phys. **24**, 20012 (2022).

[151] B. I. Schneider, T. N. Rescigno, B. H. Lengsfield, and C. W. McCurdy, Accurate *ab initio* treatment of low-energy electron collisions with ethylene, Phys. Rev. Lett. **66**, 2728 (1991).

[152] V. Zhaunerchyk, M. Kamińska, M. Mucke, R. J. Squibb, J. H. D. Eland, M. N. Piancastelli, L. J. Frasinski, J. Grilj, M. Koch, B. K. McFarland, E. Sistrunk, M. Gühr, R. N. Coffee, C. Bostedt, J. D. Bozek, P. Salén, P v D. Meulen, P. Linusson, R. D. Thomas, M. Larsson *et al.*, Disentangling formation of multiple-core holes in aminophenol molecules exposed to bright X-FEL radiation, J. Phys. B: At., Mol. Opt. Phys. **48**, 244003 (2015).

[153] M. Mucke, V. Zhaunerchyk, L. J. Frasinski, R. J. Squibb, M. Siano, J. H. D. Eland, P. Linusson, P. Salén, P. Meulen, R. D. Thomas *et al.*, Covariance mapping of two-photon double core hole states in C_2H_2 and C_2H_6 produced by an x-ray free electron laser, New J. Phys. **17**, 073002 (2015).

[154] P. Lablanquie, F. Penent, and Y. Hikosaka, Multi-electron co-incidence spectroscopy: Double photoionization from molecular inner-shell orbitals, J. Phys. B: At., Mol. Opt. Phys. **49**, 182002 (2016).

[155] M. L. Murillo-Sánchez, S. Marggi Poullain, J. J. Bajo, M. E. Corrales, J. González-Vázquez, I. R. Solá, and L. Bañares, Halogen-atom effect on the ultrafast photodissociation dynamics of the dihalomethanes CH2ICl and CH2BrI, Phys. Chem. Chem. Phys. **20**, 20766 (2018).

[156] J. W. L. Lee, D. S. Tikhonov, P. Chopra, S. Maclot, A. L. Steber, S. Gruet, F. Allum, R. Boll, X. Cheng, S. Düsterer *et al.*, Time-resolved relaxation and fragmentation of polycyclic aromatic hydrocarbons investigated in the ultrafast XUV-IR regime, Nat. Commun. **12**, 6107 (2021).

[157] F. Allum, V. Music, L. Inhester, R. Boll, B. Erk, P. Schmidt, T. M. Baumann, G. Brenner, M. Burt, P. V. Demekhin *et al.*, A localized view on molecular dissociation via electron-ion partial covariance, Commun. Chem. **5**, 42 (2022).

[158] A. J. Howard, M. Britton, Z. L. Streeter, C. Cheng, R. Forbes, J. L. Reynolds, F. Allum, G. A. McCracken, I. Gabalski, R. R. Lucchese, C. W. McCurdy, T. Weinacht, and P. H. Bucksbaum, Filming enhanced ionization in an ultrafast triatomic slingshot, Commun. Chem. **6**, 81 (2023).

[159] D. Rolles, Time-resolved experiments on gas-phase atoms and molecules with XUV and x-ray free-electron lasers, Adv. Phys.: X **8**, 2132182 (2023).

[160] T. Helgaker, S. Coriani, P. Jørgensen, K. Kristensen, J. Olsen, and K. Ruud, Recent advances in wave function-based methods of molecular-property calculations, Chem. Rev. **112**, 543 (2012).

[161] A. Szabo, A. Szabó, and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory* (Macmillan, New York, 1982).