The Journal of Chemical Physics

RESEARCH ARTICLE | SEPTEMBER 25 2023

Quantitative performance analysis and comparison of optimal-continuum Gaussian basis sets for high-harmonic generation spectra

C. Morassut ; E. Coccia ; E. Luppi

J. Chem. Phys. 159, 124108 (2023)
https://doi.org/10.1063/5.0153825

The Journal of Chemical Physics ARTICLE pubs.aip.org/aip/jcp

Quantitative performance analysis and comparison of optimal-continuum Gaussian basis sets for high-harmonic generation spectra

Cite as: J. Chem. Phys. 159, 124108 (2023); doi: 10.1063/5.0153825
Submitted: 12 April 2023 • Accepted: 24 August 2023 •
Published Online: 25 September 2023

C. Morassut,[1,2] E. Coccia,[2,a)] and E. Luppi[1,a)]

AFFILIATIONS

[1] Laboratoire de Chimie Théorique, Sorbonne Université, CNRS, Paris F-75005, France
[2] Dipartimento di Scienze Chimiche e Farmaceutiche, Università di Trieste, Via L. Giorgieri 1, 34127 Trieste, Italy

[a)] **Authors to whom correspondence should be addressed:** ecoccia.units.it and eleonora.luppi@sorbonne-universite.fr

ABSTRACT

Quantum-chemistry methods in the time domain with Gaussian basis sets are increasingly used to compute high-harmonic generation (HHG) spectra of atomic and molecular systems. The quality of these approaches is limited by the accuracy of Gaussian basis sets to describe continuum energy states. In the literature, optimal-continuum Gaussian basis sets have been proposed: Kaufmann *et al.* [J. Phys. B: At., Mol. Opt. Phys. **22**, 2223 (1989)], Woźniak *et al.* [J. Chem. Phys. **154**, 094111 (2021)], Nestmann and Peyerimhoff [J. Phys. B: At., Mol. Opt. Phys. **23**, L773 (1990)], Faure *et al.* [Comput. Phys. Commun. **144**, 224 (2002)], and Krause *et al.* [J. Chem. Phys. **140**, 174113 (2014)]. In this work, we have compared the performances of these basis sets to simulate HHG spectra of H atom at different laser intensities. We have also investigated different strategies to balance basis sets with these continuum functions, together with the role of angular momentum. To quantify the performance of the different basis sets, we introduce local and global HHG descriptors. Comparisons with the grid and exact calculations are also provided.

I. INTRODUCTION

High-harmonic generation (HHG) is a highly nonlinear optical phenomenon that provides light with an attosecond (10^{-18} s) duration, which is then used for investigating ultrafast phenomena in time-resolved spectroscopies.[1–11] The HHG signal itself encodes information on the electronic structure and dynamics of the target. Investigating HHG signals leads to HHG spectroscopy, which is applied to atoms and molecules,[3,12–32] solids,[33] and liquids.[34]

HHG spectra typically cover a broad spectral range, from visible light to soft x-rays, and have a particular shape: a rapid decrease of the intensity for the low-order harmonics, then a plateau region where the harmonics have almost constant intensity, and finally, a rapid cutoff, beyond which no harmonics are observed.

The physical nature of HHG has been explained by the semiclassical three-step model (3SM).[35,36] In the 3SM, an electron interacts with the laser and escapes from the nucleus through tunnel ionization. Then, it is accelerated in the continuum by the electric field until the latter changes its sign. Therefore, the electron is re-accelerated back to the nucleus, where it recombines, emitting a photon with a frequency equal to an integer multiple of the frequency of the laser. From this model, one can also obtain the maximum energy the field can provide to the electron, known as the cutoff energy ($E_{\text{cutoff}} = I_p + 3.17U_p$), where $U_p = E_0^2/(4\omega_0)^2$ is the classic ponderomotive energy, in which E_0 is the maximum amplitude of the pulse and ω_0 is the frequency of the pulse.

The 3SM has been successfully used to capture the physics behind the HHG process. However, the 3SM cannot describe the role of electron correlation, interference between molecular orbitals (MOs), and bound states in the ionization and recombination processes. Therefore, a more sophisticated approach is needed to capture the complexity of the electronic structure and dynamics under the influence of a strong laser-field.

For this reason, quantum-chemistry methods, typically in Gaussian basis sets, became popular in recent years[37] in the attosecond research field and particularly in the simulation of HHG spectroscopy for atomic and molecular systems.[37–42]

The merit of Gaussian functions is that they are computationally advantageous with respect to other basis sets, such as spatial grids[12,43–51] and B-splines.[52–54] However, at the present stage, the main limitation of Gaussian basis sets is the description of continuum energy states.[55] This issue depends on the localized nature of Gaussian functions, which makes describing states over large distances and with rapid oscillations, as in the case of continuum states, extremely challenging.

The search for efficient Gaussian basis sets able to fix the continuum is not limited to HHG spectroscopy; it is also extendable to other physical processes involving electronic transitions in the continuum such as photoionization[52,56,57] or above-threshold ionization.[55]

In order to overcome this issue, different recipes have been reported in the literature to define Gaussian functions optimal for the continuum.[58–65] In the work of Kaufmann *et al.*,[58] Gaussian-type orbitals (GTOs) are constructed with exponents fitted to maximize their overlap with Slater-type orbitals (STOs). Recently, Woźniak *et al.*[59] went beyond the approach proposed by Kaufmann *et al.*[58] by proposing a scheme relying on simultaneous fitting, in a predefined energy range, multiple Gaussian functions to a sequence of Slater functions in order to find an approximated global minimum. The core element of this method is the definition of the energy range of the states that are energetically accessible by the electrons; therefore, they proposed the name active range-optimized (ARO) basis set.

A different approach to obtaining optimal-continuum Gaussians was suggested in the context of electron–molecule scattering by Nestmann and Peyerimhoff.[60] Exponents are indeed obtained by fitting a set of spherical Bessel functions, i.e., the spherically adapted continuum eigenfunctions for zero potential, with a linear combination of GTOs. The method has been further extended by Faure *et al.*[61] by fitting Coulomb functions with GTOs.

Gaussian basis sets to describe strong-field ionization were also proposed by Krause *et al.* in Ref. 62. In this case, the basis set is built starting from an even tempered Gaussian basis set with a sufficient number of diffuse functions interacting with a complex absorbing potential, necessary for the simulation of the ionization process.[62] In this basis set, the exponents have not been specifically fitted to a continuum function,[58–61] but the generating procedure permits us to consider this basis set as an optimal-continuum Gaussian basis set as well.

Another strategy to improve the description of the continuum has been to use complex Gaussian functions, which naturally describe the oscillating behavior of the continuum wave function.[63–65] Gaussians with complex exponents have an intrinsic oscillating behavior, which renders them appropriate to represent continuum functions, even in the high-energy range. These basis sets have been used to describe atomic and molecular photoionization.[63–65]

In this work, we aim to compare the performances of different real optimal-continuum Gaussian basis sets for calculating the HHG spectra in atomic hydrogen at various pulse intensities within the real-time time-dependent configuration interaction (RT-TD-CI) framework. Therefore, we focus on how the continuum is modeled, how a Gaussian basis set is balanced in terms of core, valence, diffuse, and continuum functions, and the role of angular momentum in Gaussian-based calculations for HHG spectra. In fact, using Gaussians optimized for the continuum is not yet a standard procedure, as it is instead for Gaussian calculations involving bound states. Therefore, it is important to establish a robust computational protocol indicating how to construct Gaussian basis sets to correctly describe electron dynamics in the continuum. Obtaining this type of protocol is a mandatory step to approach HHG spectroscopy in atoms and molecules and, consequently, extract meaningful physics from experiments.

Some of the authors already have proposed computational strategies to tackle the problem of Gaussians for continuum states of atomic and molecular systems[14,25,31,55,66–69] based on adding optimal-continuum Gaussian basis sets, i.e., with optimized exponents for the continuum, on top of a diffuse standard Gaussian basis set for bound states.[14,25,31,55,66–69] The exponents optimized for the continuum were taken from Kaufmann *et al.*[58] Here, we provide a systematic comparison among several Gaussian basis sets built up with a similar strategy that exploits the work of Kaufmann *et al.*,[58] Woźniak *et al.*,[59] Nestmann and Peyerimhoff[60] together with Faure *et al.*,[61] and Krause *et al.*[62] To quantify the performance of the different basis sets over the description of HHG spectra, we introduce local and global HHG descriptors. We also provide comparisons with exact and grid calculations.

The article is organized as follows: in Sec. II, the theoretical framework is given. Then, in Sec. III, the computational details are collected, and in Sec. IV, the results are shown and discussed. Conclusions are in Sec. V.

II. THEORY

A. Real-time time-dependent configuration-interaction (RT-TD-CI)

The time-dependent Schrödinger equation in atomic units for the H atom in an external time-dependent electric-field [$\mathbf{E}(t)$] in the length gauge is

$$i\frac{\partial|\Psi(t)\rangle}{\partial t} = (\hat{H}_0 + \hat{V}(t))|\Psi(t)\rangle, \tag{1}$$

where $H_0(\mathbf{r}) = -\nabla^2/2 - 1/|\mathbf{r}|$ is the time-independent field-free Hamiltonian and $V(\mathbf{r},t) = \mathbf{r}\cdot\mathbf{E}(t)$ is the interaction potential between the atom and the external time-dependent electric-field in the dipole approximation. We consider the case of an electric field $\mathbf{E}(t)$ linearly polarized along the z-axis, representing a laser pulse,

$$\mathbf{E}(t) = E_0\mathbf{n}_z\,\sin(\omega_0 t + \phi)f(t), \tag{2}$$

where E_0 is the maximum field strength, $\mathbf{n}_z$ is a unit vector along the z axis, ω_0 is the pulse frequency, ϕ is the phase, and $f(t)$ is a trapezoidal envelope function calculated as

$$f(t) = \begin{cases} t/T_0, & 0 \le t < T_0, \\ 1, & T_0 \le t < (oc-1)\times T_0, \\ oc - t/T_0, & (oc-1)\times T_0 \le t < oc\times T_0, \end{cases} \tag{3}$$

with $T_0 = 2\pi/\omega_0$ and oc as an integer, which indicates the number of optical cycles. The pulse duration is $oc \times T_0$.

The time-dependent Schrödinger equation [Eq. (1)] is solved by representing $|\Psi(t)\rangle$ in RT-TD-CI as

$$|\Psi(t)\rangle = \sum_{k\geq 0} c_k(t)|\Psi_k\rangle, \tag{4}$$

composed of the Hartree–Fock ground state $|\Psi_0\rangle$ for $k = 0$ and of the field-free CI excited states $|\Psi_k\rangle$ for $k > 0$. In applying this many-electron methodology to our specific case of the H atom, considering a given basis set of size K, we obtain one occupied molecular orbital $|\phi_1\rangle$ and $v = K - 1$ virtual orbital $|\phi_a\rangle$. Therefore, $|\Psi_0\rangle$ coincides with $|\phi_1\rangle$ and the excited states are single-orbital substitutions $|\phi_a\rangle$.

Consequently, for the H atom, $|\Psi(t)\rangle$ is

$$|\Psi(t)\rangle = c_0(t)|\phi_1\rangle + \sum_{k=2}^{K} c_k(t)|\phi_k\rangle. \tag{5}$$

Inserting the time-dependent wavefunction, Eq. (5), into Eq. (1) and projecting with bras corresponding to the Hartree–Fock ground state and the CI excited states, we obtain a time-dependent equation for the coefficients,

$$i\frac{dc_k(t)}{dt} = \sum_s (E_k\delta_{ks} + \mathbf{r}_{ks}\cdot\mathbf{E}(t))c_s(t), \tag{6}$$

where E_k represents the Hartree Fock energy ($k = 0$) and CI excitation energies ($k > 0$), $\mathbf{r}_{ks} = \langle\Psi_k|\hat{\mathbf{r}}|\Psi_s\rangle$, and the initial condition at $t = 0$ is chosen to be the ground state represented in the chosen basis.

In order to solve Eq. (6), we approximate the time integral by discretizing the time of the propagation and by making use of a split propagator technique, as the Hamiltonian is the sum of two independent terms:[14] the field-free Hamiltonian $\hat{H}_0$ and the interaction term $\hat{V}(t)$. From the knowledge of time-dependent wavefunctions, the induced time-dependent position is written as

$$\mathbf{r}(t) = \sum_{sk} c_s^*(t)c_k(t)\langle\Psi_s|\hat{\mathbf{r}}|\Psi_k\rangle, \tag{7}$$

whose Fourier transform gives the power spectrum of the HHG,

$$P(\omega) = \left|\frac{1}{t_f - t_i}\int_{t_i}^{t_f} W(t)\mathbf{r}(t)\cdot\mathbf{n}_z\, e^{-i\omega t}dt\right|^2, \tag{8}$$

where $W(t)$ is the Hanning window function and t_i and t_f are, respectively, the initial and final times of the propagation.

B. Ionization model

To correct the incompleteness of the Gaussian basis sets under study, we used the modified version of the heuristic lifetime model of Klinkusch *et al.*[70] that we presented in Ref. 14 and which we call the double-*d* heuristic lifetime model (DdHLM).

In the DdHLM, the field-free eigenstates $|\Psi_k\rangle$ above the ionization threshold are interpreted as non-stationary states. Therefore, in time propagation, the energy E_k is replaced by the complex quantity $E_k - \frac{i}{2}\Gamma_k$, where Γ_k is the inverse lifetime of state k, defined as follows:

$$\Gamma_k = \begin{cases} 0 & \text{if } E_k < I_p, \\ \sqrt{2E_k}/d_0 & \text{if } I_p \leq E_k \leq E_{\text{cutoff}}, \\ \sqrt{2E_k}/d_1 & \text{if } E_k > E_{\text{cutoff}}, \end{cases} \tag{9}$$

where I_p is the ionization potential and E_{cutoff} is the 3SM cutoff energy.

The DdHLM contains two different empirical parameters (d_0 and d_1), which makes this model more flexible than the original approach, which instead contains only one parameter. The parameter d_0 is usually larger than d_1 because Γ_k for states belonging to $I_p \leq E_k \leq E_{\text{cutoff}}$ has to be smaller than Γ_k for states belonging to $E_k \geq E_{\text{cutoff}}$. This reflects the physical motion of the electron in the continuum and its ionization probability.

In the DdHLM, we have observed that the results are mainly affected by the choice of d_0. This is because d_0 affects those states that are mainly involved in the HHG process based on the physical interpretation of the 3SM. Following the 3SM, d_0 can be interpreted as the maximum electron excursion after ionization, i.e., $d_0 = E_0/\omega_0^2$.

A similar model has been recently proposed in Ref. 71, where the definition of Γ_k generally, i.e., for systems with more than one electron, depends on the virtual orbital energy rather than on the state one.

C. Local and global descriptors for HHG

An accurate theoretical description of HHG depends both on the ability of the basis set to properly describe the continuum energy states and on the choice of the empirical parameters d_0 and d_1 of the DdHLM. Therefore, to appreciate the correctness of a basis set, it is also necessary to take into account the strategy to overcome its incompleteness.

As a measure of the correctness of the Gaussian basis set, we defined local and global descriptors that permit comparing Gaussian-based spectra with a reference calculation.

The local descriptor is

$$\delta_i = h_i^{\text{Gaussian}} - h_i^{\text{Reference}}, \tag{10}$$

where i is an index indicating the ith harmonic, and h_i^{Gaussian} ($h_i^{\text{Reference}}$) is the harmonic peak value calculated using the Gaussian (reference) representation corresponding to the ith harmonic. In our calculations, the reference is the grid calculation. The value of h_i^{X}, with X = Gaussian or Reference, is taken as peak intensity exactly at the harmonic value, as shown in Fig. 1.

The δ_i descriptor provides an estimation of the HHG accuracy as a function of the harmonic order, possibly showing deviations with respect to the reference spectrum.

The definition of a global descriptor is more delicate. In fact, now we rely on only one parameter that has to be able to capture the most important features of the basis set. For this reason, in this case, we focused only on the harmonic peaks and defined the different global descriptors,

$$\Delta_{\text{N}} = \sum_{i=1}^{N}\frac{|\delta_i|}{N}, \tag{11}$$

and

$$\Upsilon_N = \sum_{i=1}^{N}\frac{\delta_i^2}{N}, \tag{12}$$

where N is the number of harmonics considered in the estimation of Δ (Υ). Both descriptors provide information about the average HHG behavior in a certain energy range defined by the choice of N.

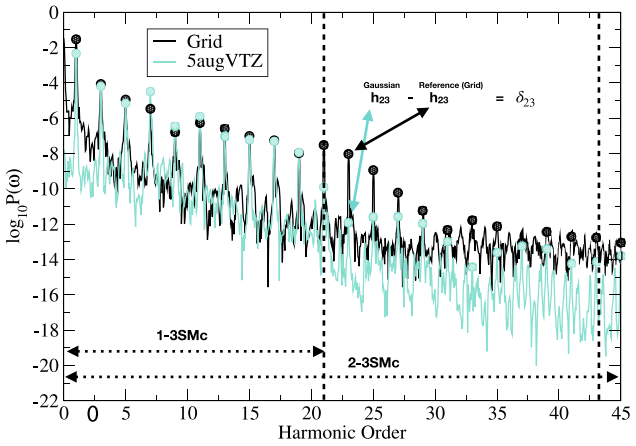

FIG. 1. Representation of the calculation of the δ_i [see Eq. (10)] for a given HHG spectrum: the difference between the maximum of the *i*th peak of the spectrum calculated with Gaussian (colored line) and the reference spectrum (black line) is performed. Here, we show an example for the 23rd harmonic. We also represent the 1-3SMc, which runs up to the 21st harmonic, and the 2-3SMc, which runs up to the 43rd harmonic (see below). A reference HHG spectrum has been computed in a spatial grid (see Sec. III).

Moreover, Woźniak *et al.*[59] proposed a different global descriptor, which is calculated as a correlation distance between two HHG signals. This descriptor is defined by considering the full frequency dependent HHG spectrum and not only the harmonic peaks. Therefore, it is not directly comparable with the descriptors we propose here. However, inspired by the work of Woźniak *et al.*,[59] we calculated a variant of their descriptor, which is

$$\check{D}_{\mathrm{corr,N}} = \sum_{i=1}^{N} \left(h_i^{\mathrm{Reference}} \cdot h_i^{\mathrm{Reference}} - h_i^{\mathrm{Reference}} \cdot h_i^{\mathrm{Gaussian}}\right). \qquad (13)$$

III. COMPUTATIONAL DETAILS

Here, we propose different strategies to obtain a balanced basis set for bound and continuum energy states. A general recipe about how to construct a Gaussian basis set capable of properly describe, at the same time, the bound and continuum properties of a system has not yet been proposed in the literature. The main idea explored is to enrich a standard Dunning basis set[72] by adding a number of diffuse and continuum functions. Then, we focus on how to create a basis set able to properly reproduce the continuum properties of the hydrogen atom in a balanced manner and also on the role of angular momentum.

We start by considering the Gaussian Dunning basis set cc-pVTZ augmented with five diffuse functions for the description of Rydberg states,[73] obtaining the 5-aug-cc-pVTZ basis set (5augVTZ short label). Then, we have customized this basis set by adding optimal-Gaussian continuum functions[58–62] at each angular momentum (*s*, *p*, and *d*). We have added to the 5augVTZ eight *s*, seven *p*, and six *d* optimal-continuum Gaussian exponents, which corresponds to double the number of functions (core, valence, and diffuse) for each angular momentum in the 5augVTZ. We refer to this strategy as 8-7-6.

Following this protocol, we have constructed the optimal-continuum Gaussian basis sets: 5augVTZK,[58] 5augVTZARO30,[59] 5augVTZARO90,[59] and 5augVTZB12.[60,61] The numbers 30 and 90, respectively, in the ARO30 and ARO90 basis sets[59] indicate the maximum STO principal quantum number used for the fitting. Instead, the label B12 was chosen because a boundary radius of $12a_0$ was used for the construction of the GTOs. A special construction was performed for the Krause *et al.*[62] basis set (S). Since in this case the basis set is built starting from an even tempered Gaussian basis set, there is no need to use the 5augVTZ on the bottom of S. We only pay attention to considering the same number of exponents for each angular momentum, identical to the number we used for the optimal-continuum basis sets.

It is important to point out that the strategy 8-7-6 is not unique. For example, in previous studies by some of us,[14,68] we did not use the 8-7-6 strategy, but we just added to each angular momentum a number of exponents able to reproduce a reasonable density of the continuum. The reasonableness is clearly related to the intensity of the laser used in the simulation. Therefore, we have also explored a different strategy to balance the basis set. We constructed two sets of basis functions, adding on top of the 5augVTZ five and seven continuum functions for each angular momentum, respectively. We refer to this strategy, respectively, as 5-5-5 and 7-7-7. We have investigated this choice for the K[14,58] and ARO90[59] basis sets, obtaining the 5augVTZ5K and the 5augVTZ5A90 (5-5-5) and the 5augVTZ7K and the 5augVTZ7A90 (7-7-7). In cases 8-7-6 and 7-7-7, we can also study the effect of a different distribution of functions on the angular momenta for HHG spectra, with the total number of functions remaining unchanged.

Moreover, we have also investigated another series of basis sets in order to explore the role of angular momentum in continuum functions. We have considered the Gaussian Dunning basis set cc-pV5Z[72] augmented with five diffuse functions for the description of Rydberg states,[73] obtaining the 5aug-cc-pV5Z basis set (5augV5Z short label). Then, also in this case, we have customized this basis set by adding optimal-Gaussian continuum functions[58–62] at each angular momentum (*s*, *p*, *d*, *f*, and *g*). The same strategy as for 8-7-6 basis sets has been applied, resulting in ten *s*, nine *p*, eight *d*, seven *f*, and six *g* optimal-continuum Gaussian exponents (10-9-8-7-6 basis sets). Following this protocol, we constructed the optimal-continuum Gaussian basis sets: 5augV5ZK,[58] 5augV5ZARO30,[59] 5augV5ZARO90,[59] and 5augV5ZB14.[60,61] The label B14 was chosen because a boundary radius of $14a_0$ was used for the construction of the GTOs. In addition, in this case, we constructed the S** basis set[62] considering the same number of exponents for each angular momentum, identical to the number we used for the optimal-continuum basis sets. We used the particular label with two asterisks to indicate that S** has two more angular momenta than S.

All the basis sets are reported in the supplementary material (see Tables I–XIV).

We have performed a configuration interaction calculation using the QChem software package[74] to obtain the electronic structure of the ground and excited states of hydrogen with the above-mentioned different basis sets. RT-TD-CI simulations have then been carried out by employing the Light code,[31,32,66,73,75] which propagates the electronic wavepacket under the influence of a time-dependent laser-field and in the presence of an ionization model.

We have set the laser pulse wavelength to 800 nm (1.55 eV/0.057 a.u.), while the pulse intensity ($I = \frac{|E_0|^2}{2}$) is equal to $I_1 = 5 \times 10^{13}$ W cm^{-2}, $I_2 = 1 \times 10^{14}$ W cm^{-2}, and $I_3 = 2 \times 10^{14}$ W cm^{-2}. The time step is 0.0538 a.u., and the pulse duration is equal to 20 optical cycles, i.e., 53 fs. The phase ϕ has been set to zero.

We have also performed HHG calculations in the length gauge on a radial grid.[14,55] The same pulse parameters and time step as for Gaussian-based calculations have been used. For $I_1 = 5 \times 10^{13}$ W cm^{-2}, we used a box of $r_{max} = 1024\ a_0$ with $l_{max} = 127$; for $I_2 = 1 \times 10^{14}$ W cm^{-2}, we used a box of $r_{max} = 4096\ a_0$ and $l_{max} = 255$, and for $I_3 = 2 \times 10^{14}$ W cm^{-2}, we used a box of $r_{max} = 4096\ a_0$ and $l_{max} = 511$. A mask function has been applied to account for ionization. The chosen mask function is equal to $\cos^{1/8}$.[76] The convergence of the grid spectrum with respect to the number of angular momenta is reported in Figs. 1 and 2 of the supplementary material. Convergence with alternative gauges, such as velocity and acceleration ones, has been explored and reported elsewhere.[14,50,55]

IV. RESULTS AND DISCUSSION

A. Modeling the continuum with the 8-7-6 basis sets

In Table I, we report the total number of energy states for the Gaussian basis sets with the 5augVTZ and the 8-7-6 strategy for the continuum, together with the corresponding number of bound (below zero) and continuum (above zero) ones. In order to quantitatively compare the results, we set the maximum energy E^{max} at around 6 Ha, neglecting sparse states with too high positive energy. The precise value for each basis set is reported in Table I.

In Fig. 2, the energies E_k are plotted as a function of the number of states. The 5augVTZ describes up to 3.454 Ha, and it is composed of 33 bound states and 26 continuum states (Table I). Adding to the 5augVTZ the optimal-continuum exponents, the number of bound states does not change for the 5augVTZARO30, while it slightly increases for the 5augVTZK, the 5augVTZARO90, the 5augVTZB12, and the S. Instead, the number of continuum states strongly increases for all the basis sets. From Table I and Fig. 2, we observe that all the optimal-continuum Gaussian basis sets have a similar behavior.

We analyzed the oscillatory behavior of the radial wavefunctions for positive energies. In Fig. 3, we report the p radial wavefunction at a given positive energy (provided by the quantum-chemistry calculation and reported in Table XV of the supplementary material)

TABLE I. Number of total, bound, and continuum energy states for each basis set with the 5augVTZ and the 8-7-6 5augVTZ-based. The highest energy value described by each basis set, E^{max}, is also reported. These basis sets have s, p, and d angular momenta.

Basis set	Total	Bound	Continuum	E^{max} (Ha)
5augVTZ	59	33	26	3.454
5augVTZK	104	37	67	6.396
5augVTZARO30	96	33	63	6.331
5augVTZARO90	104	42	62	5.921
5augVTZB12	103	36	67	5.999
S	95	38	57	6.646

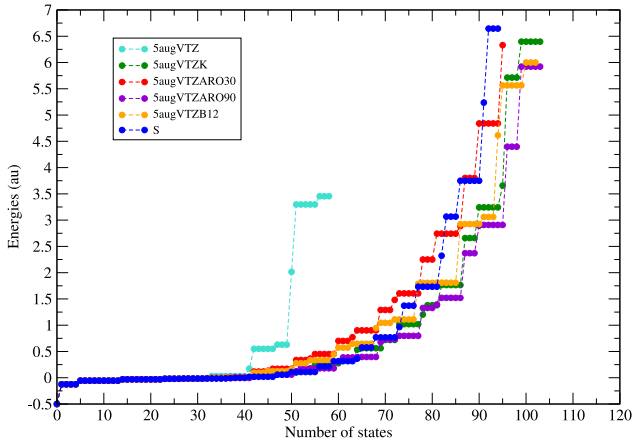

FIG. 2. Representation of the energy of states for the 8-7-6 5augVTZ-based Gaussian basis sets.

for each Gaussian basis set. An analytical solution at that energy is also reported for comparison. With the exception of 5augVTZ and S, which do not contain any specific Gaussian function for the continuum, all the other basis sets perform well in reproducing the oscillatory behavior of a continuum wavefunction. Due to the localized nature of Gaussians, such oscillations cannot be maintained at relatively large distances, which tend to disappear beyond around $13a_0$. The s and d radial wavefunctions for continuum energies are reported, respectively, in Figs. 3 and 4 of the supplementary material, providing the same general trend.

Once the role of continuum-optimal Gaussian functions has been verified to be mandatory to capture the behavior of an ejected electron, we discuss here the global and local HHG descriptors of Eqs. (10)–(13). Crucial to this discussion is the role of the lifetime model in the calculation of the HHG spectra. We used the DdHLM, which implies the choice of two empirical parameters, d_0 and d_1. We found that HHG behavior is particularly sensitive to the choice of d_0 and much less sensitive to the choice of d_1.[14,31] Therefore, we set $d_1 = 0.1a_0$ and studied the dependence of the HHG spectra and the descriptors on the values of d_0.

We started considering d_0 equal to the maximum excursion length predicted by the 3SM (d_0-3SM), and we calculated HHG spectra at different laser intensities. Then, we calculated the global and local descriptors considering up to N, i.e., the number of harmonics up to the 3SM cutoff (1-3SMc) or up to approximately twice the 3SM cutoff (2-3SMc). Therefore, in the case of the 1-3SMc, we obtained $N = 15$ ($I_1 = 5 \times 10^{13}$ W cm^{-2}), $N = 21$ ($I_2 = 1 \times 10^{14}$ W cm^{-2}), and $N = 33$ ($I_3 = 2 \times 10^{14}$ W cm^{-2}), while in the case of the 2-3SMc, we got $N = 31$ ($I_1 = 5 \times 10^{13}$ W cm^{-2}), $N = 43$ ($I_2 = 1 \times 10^{14}$ W cm^{-2}), and $N = 67$ ($I_3 = 2 \times 10^{14}$ W cm^{-2}).

However, as the choice of d_0 can significantly affect the HHG spectrum, we also took into account two other values of d_0. We considered the d_0 (d_0-best1), which minimizes Δ_N when N is taken up to 1-3SMc, and the d_0 (d_0-best2), which still minimizes Δ_N, but in this case, N is taken up to 2-3SMc. To find the values of d_0-best1 and d_0-best2, we started from d_0-3SM, and within steps of $\pm 5a_0$, we searched for the smallest Δ_N.

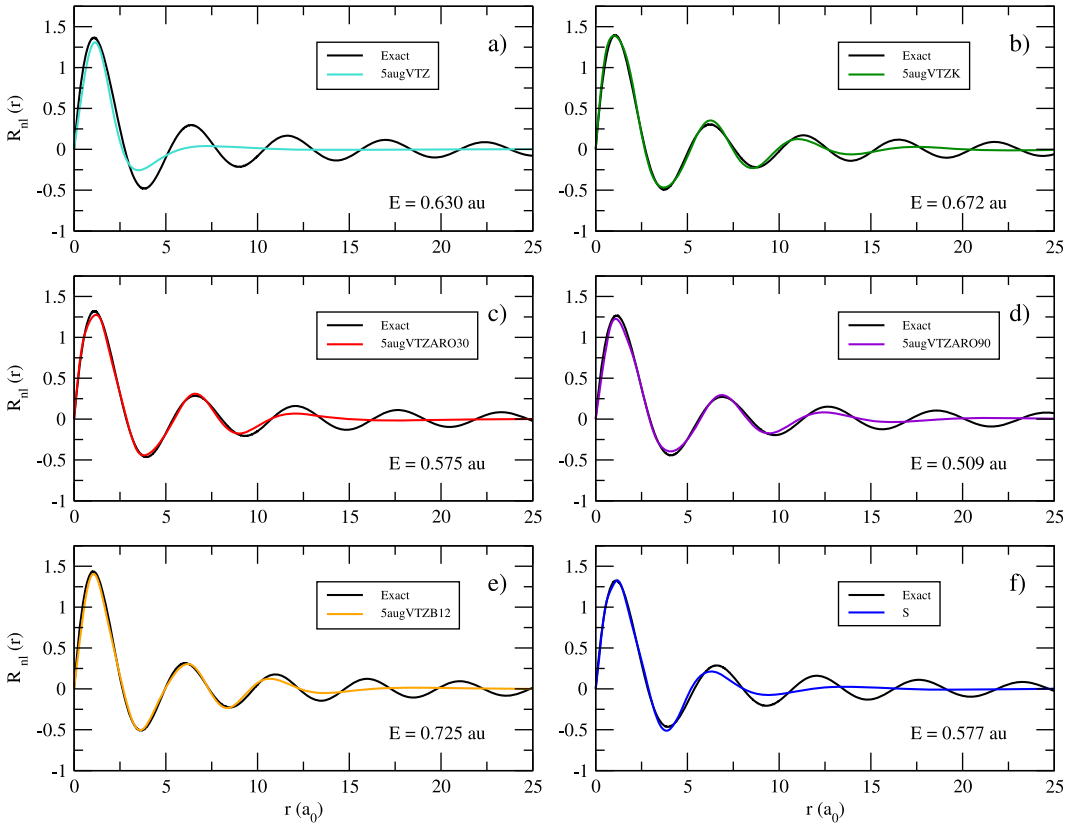

FIG. 3. Radial plot of *p* functions for the 8-7-6 5augVTZ-based Gaussian basis sets (colored line) compared with exact calculation (black line).

In Table II, we report the value of Δ_N computed with d_0-3SM and d_0-best1 for 1-3SMc at different pulse intensities. We identified for each intensity the Gaussian basis set that gives the best performance, i.e., the lowest Δ_N (bold). Taking this value as the reference, we reported in parenthesis the percentage increase[77] of the Δ_N calculated with the other basis sets. The percentage calculation has to be performed using the values of the harmonic peak intensity in the log scale. Furthermore, in Table II, we also identified

TABLE II. Global descriptor Δ_N calculated for the 5augVTZ and the 8-7-6 5augVTZ-based basis sets with d_0-3SM and d_0-best1 for 1-3SMc using the laser intensities: $I_1 = 5 \times 10^{13}$ W cm^{-2}, $I_2 = 1 \times 10^{14}$ W cm^{-2}, and $I_3 = 2 \times 10^{14}$ W cm^{-2}. For each intensity, the lowest value of global descriptor is highlighted in boldface, while the largest one is in italics.

Laser intensity		I_1		I_2		I_3	
1-3SMc		d_0 (a_0)	Δ_{15}	d_0 (a_0)	Δ_{21}	d_0 (a_0)	Δ_{33}
5augVTZ	d_0-3SM	11.65	0.485(+15%)	16.48	**0.540**	23.30	1.002(+8%)
	d_0-best1	11.65	0.485(+27%)	16.48	0.540(+2%)	20.00	0.960(+7%)
5augVTZK	d_0-3SM	11.65	0.457(+8%)	16.48	0.720(+33%)	23.30	*1.373*(+48%)
	d_0-best1	11.65	0.458(+20%)	25.00	0.674(+22%)	40.00	*1.141*(+27%)
5augVTZARO30	d_0-3SM	11.65	*0.670*(+59%)	16.48	0.832(+54%)	23.30	1.008(+9%)
	d_0-best1	35.00	*0.638*(+67%)	25.00	*0.725*(+32%)	60.00	0.991(+11%)
5augVTZARO90	d_0-3SM	11.65	0.450(+7%)	16.48	0.631(+17%)	23.30	**0.928**
	d_0-best1	11.65	0.450(+18%)	7.00	**0.551**	23.30	0.928(+4%)
5augVTZB12	d_0-3SM	11.65	0.478(+13%)	16.48	*0.919*(+70%)	23.30	0.948(+2%)
	d_0-best1	7.00	0.458(+20%)	5.00	0.602(+9%)	40.00	**0.896**
S	d_0-3SM	11.65	**0.422**	16.48	0.822(+52%)	23.30	1.156(+25%)
	d_0-best1	25.00	**0.381**	5.00	0.612(+11%)	60.00	1.122(+25%)

the Gaussian basis set that gives the worst performance, i.e., the largest Δ_N (italics).

The first general observation is that for all the Gaussian basis sets, Δ_N increases as the laser intensity increases. This is a clear indication that the performance of the basis set deteriorates when the laser is more intense. In fact, higher and higher energy continuum states are involved in the electron dynamics, and the accuracy of the basis set progressively decreases.

Considering the lowest intensity $I_1 = 5 \times 10^{13}$ W cm^{-2}, we found that the S basis set has the lowest Δ_{15} for both values of d_0. Instead, considering $I_2 = 1 \times 10^{14}$ W cm^{-2}, we found that the 5augVTZ has the lowest Δ_{21} for d_0-3SM, while the 5augVTZARO90 has the lowest Δ_{21} for d_0-best1. Finally, for $I_3 = 2 \times 10^{14}$ W cm^{-2}, we found that the 5augVTZARO90 has the lowest Δ_{33} for d_0-3SM and the 5augVTZB12 has the lowest Δ_{33} for d_0-best1.

By analyzing the results, we found that the cases where the variation is higher than 50% are due to an enhancement of only a few harmonics at an energy close to the 3SM cutoff. Despite that, overall HHG behavior continues to be reasonably accurate.

In Table III, we report the value of Δ_N calculated with the 2-3SMc. We observe, as already pointed out in Table II, that Δ_N increases as the laser intensity increases. This confirms the deterioration of the basis set's performance with the increasing the laser intensity. Moreover, for 2-3SMc, the worsening of all the basis sets seems bigger than for 1-3SMc. In fact, at a fixed laser intensity and d_0, the Δ_N is larger for 2-3SMc than for 1-3SMc. This is particularly evident for the 5augVTZ and the S basis sets. From the results in Table III, we observe that the 5augVTZARO90 and the 5augVTZB12 are the best options for all three studied intensities, while the worst ones are always the 5augVTZ and the S.

The behavior of Δ_N shows the importance of using exponents optimized for the continuum. The 5augVTZ and the S basis sets do not have a proper description of the continuum and, therefore, the deterioration of the basis set to describe the high energy states is more pronounced, as clearly shown by Table III. If one focuses on the high-energy portion and cutoff region of an HHG spectrum, using optimal Gaussian functions is mandatory.

Another important comment is that a general criterion for setting d_0 cannot be easily extracted. However, we noticed that in the case of 1-3SMc, the d_0-3SM gives a good estimation of the "best" value for d_0, even if it tends to be underestimated for I_3. In the case of 2-3SMc, d_0-3SM is still a reasonable choice for I_1 and I_2, while for I_3, d_0-best2 is much larger than d_0-3SM. The extreme case is shown for 5augVTZ, where the best value of d_0 was found by taking its infinite limit.[78] This means that the energy states have an infinite lifetime, i.e., no ionization occurs. In fact, using d_0-3SM ionization is overestimated, and too much norm of the time-dependent wavefunction is removed from the propagation and, therefore, from the HHG process. A possible solution is to exploit the incompleteness of the basis set to define an *ab initio* set of lifetimes.[67] The Δ_N values of Tables II and III are graphically reported in Fig. 5 of the supplementary material.

In order to verify the observed trend, we compared Δ_N with the other two global descriptors: Υ_N (Tables XVI and XVII of the supplementary material) and $\check{D}_{corr,N}$ (Tables XVIII and XIX of the supplementary material) from Eqs. (12) and (13), respectively. We considered only d_0-3SM since d_0-best1 and d_0-best2 have been optimized for Δ_N. The results are reported in Fig. 5 of the supplementary material. For both 1-3SMc and 2-3SMc, we obtained the same trend we found with Δ_N within small percentage fluctuations. We confirm that with Υ_N and $\check{D}_{corr,N}$ descriptors, the basis sets with the worst performance are those that do not contain exponents optimized for the continuum.

By construction, the global descriptor provides an average over low-energy and high-energy harmonics. Therefore, a piece of important complementary information to global descriptors is given by the local descriptor in Eq. (10). The local descriptor provides information about the height and the number of harmonics found up to N. The number of harmonics can vary with the chosen Gaussian basis set, and therefore, the δ_i descriptor has to be applied also in the high-

TABLE III. Global descriptor Δ_N calculated for the 5augVTZ and the 8-7-6 5augVTZ-based basis sets with d_0-3SM and d_0-best2 for 2-3SMc using the laser intensities: $I_1 = 5 \times 10^{13}$ W cm^{-2}, $I_2 = 1 \times 10^{14}$ W cm^{-2}, and $I_3 = 2 \times 10^{14}$ W cm^{-2}. For each intensity, the lowest value of global descriptor is highlighted in boldface, while the largest is in italics.

Laser intensity		I_1		I_2		I_3	
2-3SMc		d_0 (a_0)	Δ_{31}	d_0 (a_0)	Δ_{43}	d_0 (a_0)	Δ_{67}
5augVTZ	d_0-3SM	11.65	0.871(+62%)	16.48	1.060(+26%)	23.30	*2.053*(+72%)
	d_0-best2	11.65	0.871(+62%)	16.48	*1.060*(+48%)	∞	*1.445*(+93%)
5augVTZK	d_0-3SM	11.65	0.583(+9%)	16.48	0.986(+17%)	23.30	1.630(+36%)
	d_0-best2	7.00	0.561(+4%)	35.00	0.721(+1%)	300.00	0.996(+33%)
5augVTZARO30	d_0-3SM	11.65	0.762(+42%)	16.48	0.913(+8%)	23.30	1.387(+16%)
	d_0-best2	20.00	0.682(+27%)	30.00	0.777(+8%)	100.00	0.885(+18%)
5augVTZARO90	d_0-3SM	11.65	**0.537**	16.48	**0.843**	23.30	1.197(+0.1%)
	d_0-best2	11.65	**0.537**	25.00	0.784(+9%)	50.00	1.004(+34%)
5augVTZB12	d_0-3SM	11.65	0.650(+21%)	16.48	0.853(1%)	23.30	**1.195**
	d_0-best2	11.65	0.650(+21%)	60.00	**0.717**	70.00	**0.750**
S	d_0-3SM	11.65	*0.916*(+71%)	16.48	*1.741*(+107%)	23.30	1.947(+63%)
	d_0-best2	25.00	*0.902*(+68%)	250.00	0.873(+22%)	600.00	1.085(+45%)

energy portion of the HHG spectrum. In other words, thanks to δ_i, we can say that a Gaussian basis set is better than another one because it is able to reproduce high-energy peaks.

In Fig. 4, we present the local descriptor δ_i, and the corresponding HHG spectra are given in Fig. 5 for $I_1 = 5 \times 10^{13}$ W cm^{-2}. If d_0-best1 and d_0-best2 coincide with d_0-3SM, only one set of data is reported, as for 5augVTZ and 5augVTZARO90. Otherwise, two or three sets of data are reported, depending on whether d_0-best1 or/and d_0-best2 differ from d_0-3SM. For this intensity, the difference with the exact reference is rather small, and the d_0-best1 and d_0-best2 give very similar results compared to the d_0-3SM. The agreement with the reference spectrum worsens at high-energy harmonics, as expected. A similar behavior of δ_i up to H15 is found for all the basis sets: a minimum at H5–H7 and a maximum at H11. The pattern becomes strongly basis-dependent beyond H15; basis sets without a proper continuum description, i.e., 5augVTZ and S, result in large deviations from the grid reference. This information suggests that the same physics is reasonably captured by all basis sets at low energy. Two conclusions can be formulated: (i) ionization is small with I_1 and, therefore, a change in the parameter governing DdHLM does not appreciably affect the HHG spectrum; (ii) finding the best d_0 is equivalent to a self-consistent correction of the incompleteness of the basis set, which at the end gives a "good" spectrum. A similar comment can be applied to 5augVTZ and 5augVTZARO90. The behavior of 5augVTZARO30 is instead not fully understood.

Corresponding HHG spectra are collected in Fig. 5. Details discussed earlier could not be captured by a simple inspection of the spectra, which look very similar to each other. That is why we decided to introduce global and local descriptors for a fine investigation of the performance of Gaussian basis sets in simulating HHG spectra.

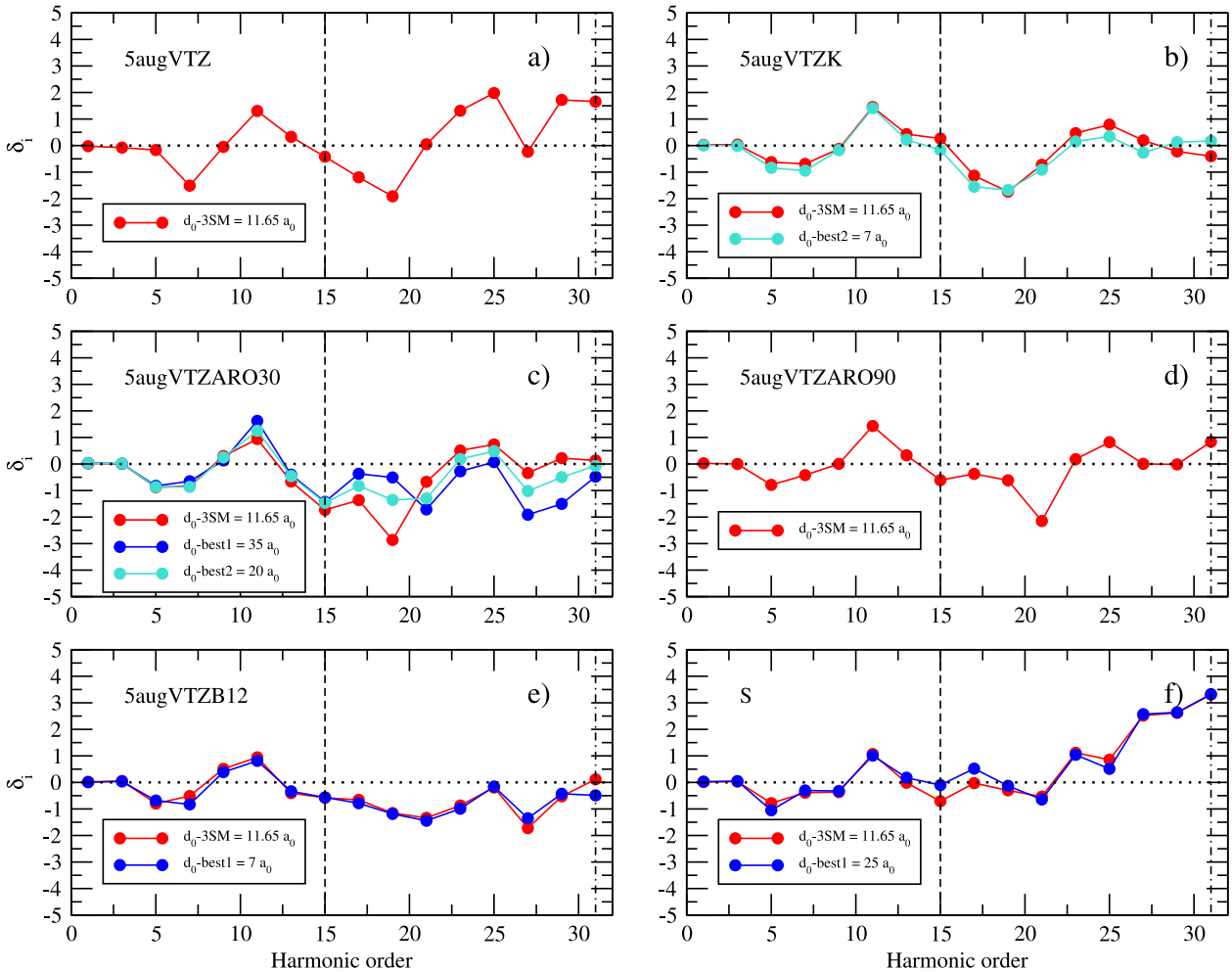

FIG. 4. Local descriptor δ_i calculated for d_0-3SM, d_0-best1, and d_0-best2. The intensity of the laser is $I_1 = 5 \times 10^{13}$ W cm^{-2}. The 8-7-6 basis sets considered are the 5augVTZ (a), the 5augVTZK (b), the 5augVTZARO30 (c), the 5augVTZARO90 (d), the 5augVTZB12 (e), and the S (f). The horizontal dotted line represents the reference grid, while the vertical lines represent the semiclassical cutoff (1-3SMc, dashed) and twice of it (2-3SMc, dotted–dashed).

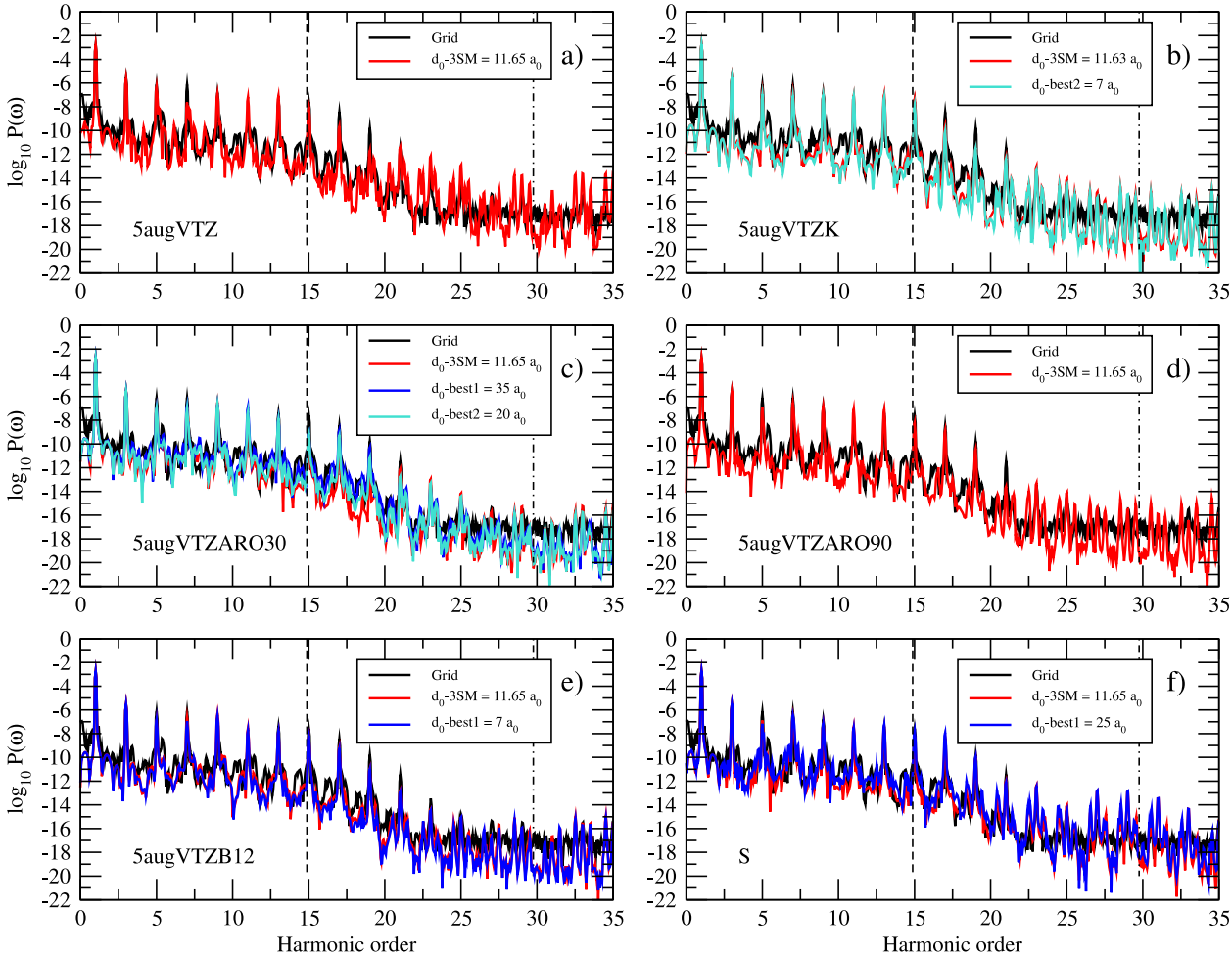

FIG. 5. HHG spectra calculated for d_0-3SM (red line), d_0-best1 (blue), and d_0-best2 (light blue), compared with grid reference (black). The intensity of the pulse considered is $I_1 = 5 \cdot 10^{13}$ W/cm^2. The basis sets considered are the 5augVTZ (a), the 8-7-6 ones: 5augVTZK (b), 5augVTZARO30 (c), 5augVTZARO90 (d), 5augVTZB12 (e), and S (f). The vertical lines represent the semiclassical cutoff (1-3SMc, dashed) and twice of it (2-3SMc, dotted–dashed).

TABLE IV. Global descriptor Δ_N calculated for 5-5-5 and 7-7-7 basis sets with d_0-3SM for 1-3SMc using the laser intensities: $I_1 = 5 \times 10^{13}$ W cm^{-2}, $I_2 = 1 \times 10^{14}$ W cm^{-2}, and $I_3 = 2 \times 10^{14}$ W cm^{-2}. Data for the corresponding 8-7-6 basis sets are also given. In parenthesis, the percentage increase (or decrease) of global descriptor is shown, taking as reference the lowest one for 8-7-6 basis set. If the reference value is referred to 5augVTZARO90 or 5augVTZK basis set, this one is also reported in boldface in this table, otherwise check bold value in Table II (for d0-3SM case).

Laser intensity		I_1		I_2		I_3	
1-3SMc		d_0 (a_0)	Δ_{15}	d_0 (a_0)	Δ_{21}	d_0 (a_0)	Δ_{33}
5augVTZK	d_0-3SM	11.65	0.457(+8%)	16.48	0.720(+33%)	23.30	*1.373*(+48%)
5augVTZ5K	d_0-3SM	11.65	0.539(+28%)	16.48	0.795(+47%)	23.30	1.012(+9%)
5augVTZ7K	d_0-3SM	11.65	0.388(-8%)	16.48	0.562(+4%)	23.30	1.053(+13%)
5augVTZARO90	d_0-3SM	11.65	0.450(+7%)	16.48	0.631(+17%)	23.30	**0.928**
5augVTZ5ARO90	d_0-3SM	11.65	0.599(+42%)	16.48	0.740(+37%)	23.30	1.251(+35%)
5augVTZ7ARO90	d_0-3SM	11.65	0.608(+44%)	16.48	0.577(+7%)	23.30	0.805(-13%)

TABLE V. Global descriptor Δ_N calculated for 5-5-5 and 7-7-7 basis sets with d_0-3SM for 2-3SMc using the laser intensities: $I_1 = 5 \times 10^{13}$ W cm^{-2}, $I_2 = 1 \times 10^{14}$ W cm^{-2}, and $I_3 = 2 \times 10^{14}$ W cm^{-2}. Data for the corresponding 8-7-6 basis sets are also reported. In parenthesis, the percentage increase (or decrease) of global descriptor is shown, taking as reference the lowest one for 8-7-6 basis set. If the reference value is referred to 5augVTZARO90 or 5augVTZK basis set, this one is also reported in boldface in this table, otherwise check the bold value in Table III (for d0-3SM case).

Laser intensity		I_1		I_2		I_3	
2-3SMc		d_0 (a_0)	Δ_{31}	d_0 (a_0)	Δ_{43}	d_0 (a_0)	Δ_{67}
5augVTZK	d_0-3SM	11.65	0.583(+9%)	16.48	0.986(+17%)	23.30	1.630(+36%)
5augVTZ5K	d_0-3SM	11.65	0.699(+30%)	16.48	0.731(-13%)	23.30	1.422(+19%)
5augVTZ7K	d_0-3SM	11.65	0.458(-17%)	16.48	0.816(-3%)	23.30	1.247(+4%)
5augVTZARO90	d_0-3SM	11.65	**0.537**	16.48	**0.843**	23.30	1.197(+0.1%)
5augVTZ5ARO90	d_0-3SM	11.65	0.640(+19%)	16.48	0.749(-11%)	23.30	1.608(+35%)
5augVTZ7ARO90	d_0-3SM	11.65	0.958(+78%)	16.48	0.560(-33%)	23.30	0.949(-21%)

The same analysis of the local descriptor δ_i for I_2 and I_3 is reported in Figs. 6 and 7 of the supplementary material, and the corresponding HHG spectra are collected in Figs. 8 and 9 of the supplementary material. We observe that with a higher intensity, deviations from the reference grid spectrum become larger, also at low energies, and differences among d_0-3SM, d_0-best1, and d_0-best2 are much more evident.

B. Balancing the basis set: 8-7-6 vs 5-5-5 and 7-7-7

Since the strategy for the construction of a balanced basis set for HHG spectra is not univocal, it is important to verify how much the conclusions formulated for the 8-7-6 are general and also applicable to different ways to define a basis set. Therefore, we compared the 5-5-5 and the 7-7-7 strategies with the 8-7-6 for the K and the ARO90 continuum functions.

In Table XX of the supplementary material, we report the total, bound, and continuum number of states for the 5augVTZ5K, 5augVTZ7K, 5augVTZ5ARO90, and 5augVTZ7ARO90 basis sets together with the 5augVTZK and the 5augVTZARO90. In addition, in this case, we considered the maximum energy $E^{\max}$ up to 6 Ha. We found that the difference in the number of states between the different strategies is not particularly significant. The same result can be observed in Fig. 10 of the supplementary material, where the energies are plotted as a function of the number of states.

The radial plots of the d wavefunctions for 5-5-5 and 7-7-7 are, respectively, plotted in Figs. 11 and 12 of the supplementary material. The results are very similar to those obtained for the 8-7-6 basis sets.

We report in Table IV, the Δ_N calculated with the 1-3SMc, and in Table V, the Δ_N calculated with the 2-3SMc. Δ_N has been computed for the three intensities I_1, I_2, and I_3 using the d_0-3SM. The corresponding HHG spectra are reported in Fig. 13 of the supplementary material.

For both 1-3SMc and 2-3SMc, we found that the 5-5-5 and 7-7-7 strategies give reasonable results with percentage variations in the same range as those obtained with the 8-7-6 strategy (see Tables II and III). However, the best performance is given by the 7-7-7 strategy, with the only exception of ARO90 with I_1. This finding suggests that adding (6–7) continuum functions to the d angular momentum in the basis set appreciably affects the spectrum, while subtracting (8–7) continuum functions to the s angular momentum in the basis set is of minor importance. This could indicate that a good Gaussian basis set should be constructed by adding the continuum functions, particularly to the higher angular momenta. Concerning the 5-5-5, we observe a strong instability of the basis as a function of the laser intensity for the 2-3SMc. We believe that this can be an indication of a poor balance between diffuse and continuum functions. It is, therefore, clear that the way to add continuum functions is not unique and needs to be further investigated.

C. Role of the angular momentum

Another important feature that characterizes a Gaussian basis set needs to be discussed, i.e., the cardinal number. For calculations with a triple-zeta basis set, three angular momenta are included, i.e., s, p, and d. In the case of molecules, an angular momentum expansion can cause a mixing of angular momenta which, instead, does not happen for the hydrogen atom, which is a single-center and single-electron system.

Some of us have already explored the role of angular momenta in Gaussian-based HHG spectra. In a previous study by Luppi and Head-Gordon,[73] the effect of angular momentum in HHG using a Gaussian basis set without continuum functions was shown. Here, to correctly describe continuum states, Gaussian functions were added to ghost atoms. For a laser intensity of I = 10^{14} W/cm^2, the following basis sets were compared: s-aug-cc-pV5Z, s-aug-cc-pV5Z plus six cc-pVDZ ghost atoms, and s-aug-cc-pV5Z plus six cc-pVTZ ghost atoms. HHG spectra were satisfactorily convergent. In another work,[14] HHG spectra from 6(9)-aug-cc-pVTZ, 6(9)-aug-cc-pVQZ, and 6(9)-aug-cc-pV5Z basis sets for the laser intensity I = 10^{14} W/cm^2 were also compared.

In order to investigate the role of angular momentum with continuum functions, we have used the 5augV5Z and defined the corresponding 5augV5ZK, 5augV5ZARO30, 5augV5ZARO90, 5augV5ZB14, and S** basis sets (see Sec. III). Properties of these basis sets in terms of total number of states, bound states, and continuum states are in Table XXI of the supplementary material; the distribution of energies is in Fig. 14 of the supplementary material; and the radial plots at positive energies are collected in Fig. 15 of the supplementary material. In particular, the precise values of the positive energies are in Table XXII of the supplementary material. Finally, we also show the HHG in Fig. 16 of the supplementary material. Increasing the angular momentum in the basis set implies an augmentation of the density continuum, which is particularly significant for 5augV5ZARO90. With the exception of 5augV5Z

TABLE VI. Global descriptor Δ_N calculated for the 5augV5Z and the 8-7-6 5augV5Z-based basis sets with d_0-3SM for 1-3SMc using the laser intensities: $I_1 = 5 \times 10^{13}$ W cm^{-2}, $I_2 = 1 \times 10^{14}$ W cm^{-2}, and $I_3 = 2 \times 10^{14}$ W cm^{-2}. For each intensity, the lowest value of global descriptor is highlighted in boldface, while the largest is in italics.

Laser intensity		I_1		I_2		I_3	
1-3SMc		d_0 (a_0)	Δ_{15}	d_0 (a_0)	Δ_{21}	d_0 (a_0)	Δ_{33}
5augV5Z	d_0-3SM	11.65	**0.218**	16.48	*0.781* (51%)	23.30	*1.567* (112%)
5augV5ZK	d_0-3SM	11.65	0.382 (75%)	16.48	**0.516**	23.30	1.023 (37%)
5augV5ZARO30	d_0-3SM	11.65	0.274 (26%)	16.48	0.706 (37%)	23.30	**0.738**
5augV5ZARO90	d_0-3SM	11.65	0.330 (51%)	16.48	0.520 (1%)	23.30	0.931 (26%)
5augV5ZB14	d_0-3SM	11.65	0.327 (50%)	16.48	0.608 (18%)	23.30	0.895 (21%)
S**	d_0-3SM	11.65	*0.504* (131%)	16.48	0.613 (19%)	23.30	0.784 (6%)

TABLE VII. Global descriptor Δ_N calculated for the 5augV5Z and the 8-7-6 5augV5Z-based basis sets with d_0-3SM for 2-3SMc using the laser intensities: $I_1 = 5 \times 10^{13}$ W cm^{-2}, $I_2 = 1 \times 10^{14}$ W cm^{-2}, and $I_3 = 2 \times 10^{14}$ W cm^{-2}. For each intensity, the lowest value of global descriptor is highlighted in boldface, while the largest is in italics.

Laser intensity		I_1		I_2		I_3	
2-3SMc		d_0 (a_0)	Δ_{31}	d_0 (a_0)	Δ_{43}	d_0 (a_0)	Δ_{67}
5augV5Z	d_0-3SM	11.65	0.453 (2%)	16.48	*1.070* (106%)	23.30	*2.625* (224%)
5augV5ZK	d_0-3SM	11.65	0.827 (86%)	16.48	**0.519**	23.30	1.391 (72%)
5augV5ZARO30	d_0-3SM	11.65	**0.444**	16.48	0.589 (13%)	23.30	1.153 (42%)
5augV5ZARO90	d_0-3SM	11.65	0.947 (113%)	16.48	0.927 (77%)	23.30	**0.810**
5augV5ZB14	d_0-3SM	11.65	0.681 (53%)	16.48	0.684 (32%)	23.30	1.358 (68%)
S**	d_0-3SM	11.65	*1.115* (151%)	16.48	0.626 (21%)	23.30	1.010 (25%)

and S**, which do not contain any specific Gaussian function for the continuum, all the other basis sets reproduce well the oscillatory behavior, showing a good and enhanced fit with an exact wavefunction.

In Tables VI and VII, we report the value of Δ_N computed with the d_0-3SM, respectively, for 1-3SMc and for 2-3SMc at different pulse intensities.

In the case of 1-3SMc in Table VI, we obtain a similar trend already discussed in Table II. Concerning the case of 2-3SMc in Table VII, we also found a behavior similar to that shown in Table III, but only for I_2 and I_3. The 5augV5ZK, 5augV5ZARO90, and 5augV5ZB14 basis sets at intensity I_1 do not perform better than the simple 5augV5Z. However, it is worth noticing that in most of the cases, Δ_N is lower for the 5augV5Z-based basis sets than for the 5augVTZ-based basis sets, which means an improvement of the basis performance with higher angular momenta. We found that the 5augV5ZARO90 basis set has the best agreement with the reference spectrum at I_3, i.e., the lowest Δ_N.

Concerning the results of the 5augV5Z-based basis sets with I_1 and 2-3SMc, we attribute the reason for this behavior to a failure of the DdHLM to handle the ionization for I_1 when the basis set is "converging" to completeness.

We have tried different combinations of d_0 and d_1, but we did not find a possible improvement (see Table XXIII in the supplementary material). Therefore, our suggestion is to use a more sophisticated approach to treat ionization, e.g., extracting information about the incompleteness of the basis set at the *ab initio* level.[67]

A further indication of the importance of the ionization model is shown in Table XXIV of the supplementary material, where the Δ_N is calculated for the 5augV5Z and the 5augV5ZARO90 with d_0-best1 and d_0-best2. By optimizing the values of d_0, the values of Δ_N decrease.

These results clearly show how a systematic way of constructing basis sets adapted to strong-field electron dynamics in strong fields is still missing.

V. CONCLUSIONS

In this work, we have compared the performance of different Gaussian basis sets to simulate HHG spectra considering the hydrogen atom as the benchmark system. We have investigated: (i) the performance of different optimal-continuum Gaussian basis sets; (ii) different strategies to include the optimal-continuum Gaussian functions in a diffuse Dunning basis set; (iii) the role of the cardinal number and, therefore, of the maximum angular momentum described by the basis set, together with the optimal-continuum functions.

In all cases, we analyzed the energies and the radial wavefunctions, from which we computed the HHG spectra. The quantitative comparison has been based on global Δ_N, Υ_N, and $\check{D}_{corr,N}$ and local δ_i descriptors, considering them as reference calculations on a grid.

The performance of the basis sets has been studied over different energy ranges: the 1-3SMc (HHG up to once the 3SM cutoff) and the 2-3SMc (HHG up to twice the 3SM cutoff). A rather general

feature of 1-3SMc and 2-3SMc results is the fact that the basis sets deteriorate their performances by increasing the laser intensity.

Concerning point (i), we considered the 5augVTZ (*s*, *p*, and *d*) and the 5augVTZK,[58] 5augVTZARO30,[59] 5augVTZARO90,[59] 5augVTZB12,[60,61] and S basis sets.[62] For each angular momentum, we added the same number of continuum functions as the number of functions (core, valence, and diffuse) in the 5augVTZ, resulting in an 8-7-6 strategy. We found that in the energy region covered by the 1-3SMc, all the basis sets, including the 5augVTZ and S, performed reasonably well. The fact that, up to the 1-3SMc, the basis sets perform rather accurately is promising for the description of strong-field electron dynamics for low-energy continuum and also in more complex systems. Instead, for higher energy such as those described in the 2-3SMc, the inclusion of exponents, especially for the continuum, is necessary, particularly with high laser intensity.

The results of point (i) are obtained with the particular strategy 8-7-6 to balance the basis sets. For this reason, point (ii) has been addressed by defining the basis sets using 5-5-5 and 7-7-7 strategies for optimal continuum functions. We found that an improvement in the results is generally obtained using the 7-7-7 strategy, which indicates that it is more efficient to add optimal-continuum functions to the higher angular momenta. However, it is worth noting that the 8-7-6 still produces results with the same accuracy as the 7-7-7. Instead, the 5-5-5 strategy shows an instability as a function of the laser intensity. This is due to the wrong balance between core, valence, diffuse, and optimal-continuum functions in the basis set.

At point (iii), the effect of increasing the cardinal number of the basis set was studied by considering the 5augV5Z basis set (*s*, *p*, *d*, *f*, and *g*), on top of which we constructed the 5augV5ZK,[58] 5augV5ZARO30,[59] 5augV5ZARO90,[59] and 5augV5ZB14[60,61] and S** basis set.[62] These basis sets have been defined using the same strategy as for point (i). The effect of increasing the cardinal number is related to increasing the number of angular momenta contained in the basis set. In this case, the improvement is particularly evident at higher laser intensity.

The analysis of the performances of the basis sets to describe HHG and therefore the electron dynamics in the process involving excitations and ionization in the continuum also pointed out the importance of how to deal with the basis set's incompleteness. We used DdHLM and showed that the basis set performance can be modeled by the choice of the parameters that govern DdHLM. However, even if we demonstrate that it is possible to obtain reasonable results by tuning the parameters in DdHLM, this must be done with care. In fact, DdHLM and similar methodology play a double role: (i) taking care of basis incompleteness and (ii) physical interpretation of ionization. Therefore, a basis set such as the 5augVTZ can give reasonable HHG results by *ad hoc* modifying the parameters in DdHLM, but other physical quantities can then be wrongly described. Recently, some of the authors[67] proposed a method to extract *ab initio* lifetimes for a specific basis set, which can be promising in order to avoid ambiguity in the choice of the empirical parameters of the DdHLM. The failure of the DdHLM is probably at the origin of the not optimized performance of 5aug5VZARO90 and 5aug5VZK for 2-3SMc at I_1 and of 5augVTZ7ARO90 at I_1.

Further understanding and optimization are needed to use Gaussian basis sets correctly to simulate strong-field electron dynamics. In particular, finding a systematic way to add optimal-continuum to a diffuse basis set is still a challenge. The complexity of the problem is clearly extended to the study of molecular systems, for which a large expansion of Gaussian basis sets with respect to the angular momenta is definitely not feasible. Indeed, in recent studies by some of us, we truncate the basis set to a double zeta, on top of which diffuse and K functions are added.[31,32] A way to compensate for the small number of angular momenta (*s*, *p*, and *d*) could be to add a larger number of optimal-continuum functions for the higher angular momenta (*p* and *d*).

SUPPLEMENTARY MATERIAL

Gaussian basis sets are employed in this work. Convergence of the grid spectra with respect to the number of angular momenta. 8-7-6 basis sets: radial *s* and *d* wavefunctions and state energies; Υ_N and $\dot{D}_{\mathrm{corr,N}}$ descriptors; δ_i and Δ_N descriptors; and HHG spectra. 5-5-5 and 7-7-7 basis sets: radial *d* wavefunctions, state energies, and HHG spectra. 5aug5Z-type basis sets: state energies and radial wavefunctions, Δ_N and HHG spectra.

ACKNOWLEDGMENTS

Financial support from ICSC—Centro Nazionale di Ricerca in High Performance Computing, Big Data and Quantum Computing—funded by the European Union – NextGenerationEU is gratefully acknowledged. The authors acknowledge the University of Trieste and Sorbonne Université for computational support.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

C. Morassut: Data curation (equal); Formal analysis (equal); Investigation (equal); Validation (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). **E. Coccia**: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Supervision (equal); Validation (equal); Writing – original draft (equal); Writing – review & editing (equal). **E. Luppi**: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Supervision (equal); Writing – original draft (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] P. B. Corkum and F. Krausz, Nat. Phys. **3**, 381 (2007).
[2] S. Baker, I. A. Walmsley, J. W. G. Tisch, and J. P. Marangos, Nat. Photonics **5**, 664 (2011).
[3] P. Peng, C. Marceau, and D. M. Villeneuve, Nat. Rev. Phys. **1**, 144 (2019).

[4]G. Sansone, Nat. Photonics **14**, 131 (2020).
[5]A. J. Uzan, H. Soifer, O. Pedatzur, A. Clergerie, S. Larroque, B. D. Bruner, B. Pons, M. Ivanov, O. Smirnova, and N. Dudovich, Nat. Photonics **14**, 188 (2020).
[6]T. Gaumnitz, A. Jain, Y. Pertot, M. Huppert, I. Jordan, F. Ardana-Lamas, and H. J. Wörner, Opt. Express **25**, 27506 (2017).
[7]C. Chappuis, D. Bresteau, T. Auguste, O. Gobert, and T. Ruchon, Phys. Rev. A **99**, 033806 (2019).
[8]F. Calegari, G. Sansone, S. Stagira, C. Vozzi, and M. Nisoli, J. Phys. B: At., Mol. Opt. Phys. **49**, 062001 (2016).
[9]M. Nisoli, P. Decleva, F. Calegari, A. Palacios, and F. Martín, Chem. Rev. **117**, 10760 (2017).
[10]R. Geneaux, H. J. B. Marroux, A. Guggenmos, D. M. Neumark, and S. R. Leone, Philos. Trans. R. Soc. A **377**, 20170463 (2019).
[11]P. M. Kraus, M. Zürch, S. K. Cushing, D. M. Neumark, and S. R. Leone, Nat. Rev. Chem. **2**, 82 (2018).
[12]S. Pabst and R. Santra, Phys. Rev. Lett. **111**, 233005 (2013).
[13]D. Faccialá, S. Pabst, B. D. Bruner, A. G. Ciriolo, S. De Silvestri, M. Devetta, M. Negro, H. Soifer, S. Stagira, N. Dudovich, and C. Vozzi, Phys. Rev. Lett. **117**, 093902 (2016).
[14]E. Coccia, B. Mussard, M. Labeye, J. Caillat, R. Taieb, J. Toulouse, and E. Luppi, Int. J. Quantum Chem. **116**, 1120 (2016).
[15]A. Gorlach, O. Neufeld, N. Rivera, O. Cohen, and I. Kaminer, Nat. Commun. **11**, 4598 (2020).
[16]P. M. Kraus and H. J. Wörner, Angew. Chem., Int. Ed. **57**, 5228 (2018).
[17]M. Negro, M. Devetta, D. Faccialá, S. De Silvestri, C. Vozzi, and S. Stagira, Faraday Discuss. **171**, 133 (2014).
[18]J. Heslar, D. Telnov, and S. Chu, Phys. Rev. A **83**, 043414 (2011).
[19]X. Chu and G. C. Groenenboom, Phys. Rev. A **85**, 053402 (2012).
[20]M. Monfared, E. Irani, and R. Sadighi-Bonabi, J. Chem. Phys. **148**, 234303 (2018).
[21]B. D. Bruner, Z. Masin, M. Negro, F. Morales, D. Brambila, M. Devetta, D. Faccialá, A. G. Harvey, M. Ivanov, Y. Mairesse, S. Patchkovskii, V. Serbinenko, H. Soifer, S. Stagira, C. Vozzi, N. Dudovich, and O. Smirnova, Faraday Discuss. **194**, 369 (2016).
[22]P. Mulholland and D. Dundas, Phys. Rev. A **97**, 043428 (2017).
[23]F. Bedurke, T. Klamroth, P. Krause, and P. Saalfrank, J. Chem. Phys. **150**, 234114 (2019).
[24]A. Gordon, F. Kärtner, N. Rohringer, and R. Santra, Phys. Rev. Lett. **96**, 223902 (2006).
[25]C. F. Pauletti, E. Coccia, and E. Luppi, J. Chem. Phys. **154**, 014101 (2021).
[26]N. Kajumba, R. Torres, J. G. Underwood, J. S. Robinson, S. Baker, J. W. G. Tisch, R. de Nalda, W. A. Bryan, R. Velotta, C. Altucci, I. Procino, I. C. E. Turcu, and J. P. Marangos, New J. Phys. **10**, 025008 (2008).
[27]H. J. Wörner, J. Bertrand, P. Hockett, P. Corkum, and D. Villeneuve, Phys. Rev. Lett. **104**, 233904 (2010).
[28]C. Vozzi, M. Negro, F. Calegari, G. Sansone, M. Nisoli, S. De Silvestri, and S. Stagira, Nat. Phys. **7**, 822 (2011).
[29]C. Hutchison, R. Ganeev, M. Castillejo, I. Lopez-Quintas, A. Zair, S. Weber, F. McGrath, Z. Abdelrahman, M. Oppermann, M. Martin, D. Lei, S. Maier, J. Tisch, and J. Marangos, Phys. Chem. Chem. Phys. **15**, 12308 (2013).
[30]I. Lopez-Quintas, M. Oujja, M. Sanz, A. Benitez-Canete, C. Hutchison, R. de Nalda, M. Martin, R. Ganeev, J. P. Marangos, and M. Castillejo, Appl. Surf. Sci. **302**, 299 (2014).
[31]E. Luppi and E. Coccia, Phys. Chem. Chem. Phys. **23**, 3729 (2021).
[32]C. Morassut, E. Luppi, and E. Coccia, Chem. Phys. **559**, 111515 (2022).
[33]J. Park, A. Subramani, S. Kim, and M. F. Ciappina, Adv. Phys.: X **7**, 2003244 (2022).
[34]O. Neufeld, Z. Nourbakhsh, N. Tancogne-Dejean, and A. Rubio, J. Chem. Theory Comput. **18**, 4117 (2022).
[35]P. B. Corkum, Phys. Rev. Lett. **71**, 1994 (1993).
[36]M. Lewenstein, P. Balcou, M. Y. Ivanov, A. L'Huillier, and P. B. Corkum, Phys. Rev. A **49**, 2117 (1994).
[37]E. Coccia and E. Luppi, J. Phys.: Condens. Matter **34**, 073001 (2022).
[38]H. B. Schlegel, S. M. Smith, and X. Li, J. Chem. Phys. **126**, 244110 (2007).
[39]J. A. Sonk, M. Caricato, and H. B. Schlegel, J. Phys. Chem. A **115**, 4678 (2011).
[40]J. A. Sonk and H. B. Schlegel, J. Phys. Chem. A **115**, 11832 (2011).
[41]P. Saalfrank, F. Bedurke, C. Heide, T. Klamroth, S. Klinkusch, P. Krause, M. Nest, and J. C. Tremblay, Adv. Quantum Chem. **81**, 15 (2020).
[42]F. Bedurke, T. Klamroth, and P. Saalfrank, Phys. Chem. Chem. Phys. **23**, 13544 (2021).
[43]T. Sato, H. Pathak, Y. Orimo, and K. L. Ishikawa, J. Chem. Phys. **148**, 051101 (2018).
[44]I. Schneider, Phys. Rev. A **55**, 3417 (1997).
[45]J. S. Prauzner-Bechcicki, K. Sacha, B. Eckhardt, and J. Zakrzewski, Phys. Rev. A **78**, 013419 (2008).
[46]L.-Y. Peng and A. F. Starace, J. Chem. Phys. **125**, 154311 (2006).
[47]A. N. Grum-Grzhimailo, B. Abeln, K. Bartschat, D. Weflen, and T. Urness, Phys. Rev. A **81**, 043408 (2010).
[48]C. Ruiz, L. Plaja, L. Roso, and A. Becker, Phys. Rev. Lett. **96**, 053001 (2006).
[49]A. Bandrauk, S. Chelkowski, D. J. Diestler, J. Manz, and K. J. Yuan, Phys. Rev. A **79**, 023403 (2009).
[50]Y.-C. Han and L. B. Madsen, Phys. Rev. A **81**, 063430 (2010).
[51]S. Pabst, A. Sytcheva, O. Geffert, and R. Santra, Phys. Rev. A **94**, 033421 (2016).
[52]C. Froese Fischer and M. Idrees, Comput. Phys. **3**, 53 (1989).
[53]H. Bachau, E. Cormier, P. Decleva, J. E. Hansen, and F. Martín, Rep. Prog. Phys. **64**, 1815 (2001).
[54]M. Ruberti, P. Decleva, and V. Averbukh, Phys. Chem. Chem. Phys. **20**, 8311 (2018).
[55]M. Labeye, F. Zapata, E. Coccia, V. Veniard, J. Toulouse, J. Caillat, R. Taieb, and E. Luppi, J. Chem. Theory Comput. **14**, 5846 (2018).
[56]F. Zapata, E. Luppi, and J. Toulouse, J. Chem. Phys. **150**, 234104 (2019).
[57]K. Schwinn, F. Zapata, A. Levitt, É. Cancès, E. Luppi, and J. Toulouse, J. Chem. Phys. **156**, 224106 (2022).
[58]K. Kaufmann, W. Baumeister, and M. Jungen, J. Phys. B: At., Mol. Opt. Phys. **22**, 2223 (1989).
[59]A. P. Woźniak, M. Lesiuk, M. Przybytek, D. K. Efimov, J. S. Prauzner-Bechcicki, M. Mandrysz, M. Ciappina, E. Pisanty, J. Zakrzewski, M. Lewenstein, and R. Moszyński, J. Chem. Phys. **154**, 094111 (2021).
[60]B. M. Nestmann and S. D. Peyerimhoff, J. Phys. B: At., Mol. Opt. Phys. **23**, L773 (1990).
[61]A. Faure, J. D. Gorfinkiel, L. A. Morgan, and J. Tennyson, Comput. Phys. Commun. **144**, 224 (2002).
[62]P. Krause, J. A. Sonk, and H. B. Schlegel, J. Chem. Phys. **140**, 174113 (2014).
[63]A. Ammar, L. U. Ancarani, and A. Leclerc, J. Comput. Chem. **42**, 2294 (2021).
[64]A. Ammar, A. Leclerc, and L. U. Ancarani, in *New Electron Correlation Methods and Their Applications, and Use of Atomic Orbitals with Exponential Asymptotes*, *Advances in Quantum Chemistry Vol. 83*, edited by M. Musial and P. E. Hoggan (Academic Press, 2021), pp. 287–304.
[65]A. Ammar, A. Leclerc, and L. U. Ancarani, J. Phys.: Conf. Ser. **1412**, 152079 (2020).
[66]E. Coccia and E. Luppi, Theor. Chem. Acc. **135**, 43 (2016).
[67]E. Coccia, R. Assaraf, E. Luppi, and J. Toulouse, J. Chem. Phys. **147**, 014106 (2017).
[68]E. Coccia and E. Luppi, Theor. Chem. Acc. **138**, 96 (2019).
[69]E. Coccia, Mol. Phys. **118**, e1769871 (2020).
[70]S. Klinkusch, P. Saalfrank, and T. Klamroth, J. Chem. Phys. **131**, 114304 (2009).
[71]A. P. Woźniak, M. Przybytek, M. Lewenstein, and R. Moszyński, J. Chem. Phys. **156**, 174106 (2022).
[72]T. H. Dunning, J. Chem. Phys. **90**, 1007 (1989).
[73]E. Luppi and M. Head-Gordon, J. Chem. Phys. **139**, 164121 (2013).
[74]E. Epifanovsky, A. T. B. Gilbert, X. Feng, J. Lee, Y. Mao, N. Mardirossian, P. Pokhilko, A. F. White, M. P. Coons, A. L. Dempwolff, Z. Gan, D. Hait, P. R. Horn, L. D. Jacobson, I. Kaliman, J. Kussmann, A. W. Lange, K. U. Lao, D. S. Levine, J. Liu, S. C. McKenzie, A. F. Morrison, K. D. Nanda, F. Plasser, D. R. Rehn, M. L. Vidal, Z.-Q. You, Y. Zhu, B. Alam, B. J. Albrecht, A. Aldossary, E. Alguire, J. H. Andersen, V. Athavale, D. Barton, K. Begam, A. Behn, N. Bellonzi, Y. A. Bernard, E. J. Berquist, H. G. A. Burton, A. Carreras, K. Carter-Fenk, R. Chakraborty, A. D.

Chien, K. D. Closser, V. Cofer-Shabica, S. Dasgupta, M. de Wergifosse, J. Deng, M. Diedenhofen, H. Do, S. Ehlert, P.-T. Fang, S. Fatehi, Q. Feng, T. Friedhoff, J. Gayvert, Q. Ge, G. Gidofalvi, M. Goldey, J. Gomes, C. E. González-Espinoza, S. Gulania, A. O. Gunina, M. W. D. Hanson-Heine, P. H. P. Harbach, A. Hauser, M. F. Herbst, M. Hernández Vera, M. Hodecker, Z. C. Holden, S. Houck, X. Huang, K. Hui, B. C. Huynh, M. Ivanov, D. Jász, H. Ji, H. Jiang, B. Kaduk, S. Kähler, K. Khistyaev, J. Kim, G. Kis, P. Klunzinger, Z. Koczor-Benda, J. H. Koh, D. Kosenkov, L. Koulias, T. Kowalczyk, C. M. Krauter, K. Kue, A. Kunitsa, T. Kus, I. Ladjánszki, A. Landau, K. V. Lawler, D. Lefrancois, S. Lehtola, R. R. Li, Y.-P. Li, J. Liang, M. Liebenthal, H.-H. Lin, Y.-S. Lin, F. Liu, K.-Y. Liu, M. Loipersberger, A. Luenser, A. Manjanath, P. Manohar, E. Mansoor, S. F. Manzer, S.-P. Mao, A. V. Marenich, T. Markovich, S. Mason, S. A. Maurer, P. F. McLaughlin, M. F. S. J. Menger, J.-M. Mewes, S. A. Mewes, P. Morgante, J. W. Mullinax, K. J. Oosterbaan, G. Paran, A. C. Paul, S. K. Paul, F. Pavošević, Z. Pei, S. Prager, E. I. Proynov, D. Rák, E. Ramos-Cordoba, B. Rana, A. E. Rask, A. Rettig, R. M. Richard, F. Rob, E. Rossomme, T. Scheele, M. Scheurer, M. Schneider, N. Sergueev, S. M. Sharada, W. Skomorowski, D. W. Small, C. J. Stein, Y.-C. Su, E. J. Sundstrom, Z. Tao, J. Thirman, G. J. Tornai, T. Tsuchimochi, N. M. Tubman, S. P. Veccham, O. Vydrov, J. Wenzel, J. Witte, A. Yamada, K. Yao, S. Yeganeh, S. R. Yost, A. Zech, I. Y. Zhang, X. Zhang, Y. Zhang, D. Zuev, A. Aspuru-Guzik, A. T. Bell, N. A. Besley, K. B. Bravaya, B. R. Brooks, D. Casanova, J.-D. Chai, S. Coriani, C. J. Cramer, G. Cserey, A. E. DePrince, R. A. DiStasio, A. Dreuw, B. D. Dunietz, T. R. Furlani, W. A. Goddard, S. Hammes-Schiffer, T. Head-Gordon, W. J. Hehre, C.-P. Hsu, T.-C. Jagau, Y. Jung, A. Klamt, J. Kong, D. S. Lambrecht, W. Liang, N. J. Mayhall, C. W. McCurdy, J. B. Neaton, C. Ochsenfeld, J. A. Parkhill, R. Peverati, V. A. Rassolov, Y. Shao, L. V. Slipchenko, T. Stauch, R. P. Steele, J. E. Subotnik, A. J. W. Thom, A. Tkatchenko, D. G. Truhlar, T. Van Voorhis, T. A. Wesolowski, K. B. Whaley, H. L. Woodcock, P. M. Zimmerman, S. Faraji, P. M. W. Gill, M. Head-Gordon, J. M. Herbert, and A. I. Krylov, J. Chem. Phys. **155**, 084801 (2021).

[75] E. Luppi and M. Head-Gordon, Mol. Phys. **110**, 909 (2012).

[76] J. L. Krause, K. J. Schafer, and K. C. Kulander, Phys. Rev. A **45**, 4998 (1992).

[77] To calculate the percentage increase or decrease: First, work out the difference (increase) between the two numbers you are comparing; then divide the increase by the original number and multiply the answer by 100.

[78] In practice we have observed that with 10^5 and larger values of d_0 Δ_N remains constant.