

PAPER • OPEN ACCESS

## Site-specific tunnel-ionization in high harmonic generation in molecules

To cite this article: Laura Rego *et al* 2020 *New J. Phys.* **22** 043012

View the [article online](#) for updates and enhancements.

You may also like

- [Classical versus quantum views of intense laser pulse propagation in gases](#)  
S A Berman, C Chandre, J Dubois et al.
- [Development of high harmonic generation spectroscopy of organic molecules and biomolecules](#)  
J P Marangos
- [Attosecond technology\(ies\) and science](#)  
Jens Biegert, Francesca Calegari, Nirit Dudovich et al.



## PAPER

## Site-specific tunnel-ionization in high harmonic generation in molecules

## OPEN ACCESS

## RECEIVED

11 November 2019

## REVISED

21 February 2020

## ACCEPTED FOR PUBLICATION

9 March 2020

## PUBLISHED

21 April 2020

Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Laura Rego<sup>1</sup> , Carlos Hernández-García<sup>1</sup> , Antonio Picón<sup>2</sup> and Luis Plaja<sup>1</sup>

<sup>1</sup> Grupo de Investigación en Aplicaciones del Láser y Fotónica, Departamento de Física Aplicada, University of Salamanca, E-37008 Salamanca, Spain

<sup>2</sup> Departamento de Química, Universidad Autónoma de Madrid, Madrid, Spain

E-mail: [laura.rego@usal.es](mailto:laura.rego@usal.es)

**Keywords:** high-harmonic spectroscopy, tunnel ionization, high harmonic generation in molecules, strong field approximation, ultrafast lasers, attosecond laser pulses, photoionization of molecules

## Abstract

We demonstrate that the standard picture of strong-field tunnel-ionization from molecules should be reformulated. The extended nature of the molecular potential implies the separation of some of the molecular sites from the edge of the ionization barrier. We show that the dependence of the tunnel probability with the distance to the barrier is translated into the ionized wavepacket, modifying substantially the high-order harmonic emission. The introduction of the dependence of tunnel ionization with the molecular site significantly improves the theoretical description of high-order harmonic generation in molecules, which is used as a cornerstone in high-harmonic spectroscopy and attosecond imaging.

## 1. Introduction

High-order harmonic generation (HHG) is an extreme non-linear process induced by intense fields. In atomic or molecular gases, it can be described as a three-step process [1, 2]: first, near the maxima of the driving field's amplitude, an electronic wavepacket is tunnel-ionized from the parent atom; in the second step, the electronic wavepacket is accelerated and, after reversal of the sign of the electric field, it is redirected to the parent ion; finally, upon recollision, the electron's kinetic energy is emitted in the form of high-frequency radiation. The harmonic spectra encode information of the target structure and dynamics, that can be disentangled using high-harmonic spectroscopy (HHS) and time resolved attosecond spectroscopy techniques [3]. These procedures have been successful in retrieving the information from the HHG spectra about molecular structure [4–6], nuclear dynamics [7], molecular orbitals [8, 9], energy dispersion in solids [10], dynamics in strongly correlated systems [11], tunneling times [12], and orbital tomography [13–15].

The study of molecular systems interacting with strong laser fields has a main tool in computationally solving the time-dependent Schrödinger equation (TDSE). The detailed physics, however, is frequently hidden by the complexity of the processes involved. To this end, approximated models are needed to describe the problem in terms of fundamental physical mechanisms. Among them, those based in the strong-field approximation (SFA) [16–18] allow to establish a link between the harmonic spectral signatures and the structural details of the radiating matter system. The SFA has successfully demonstrated to reproduce the main characteristics of the harmonic spectra in atoms [19–21], and molecules [13, 22–26], and their electron and nuclear dynamics [7, 27–31].

A key ingredient in HHS is the mapping of the molecular orbital matrix elements into the harmonic spectrum and phase, as predicted from the SFA. This relation allows for the reconstruction of the molecular geometry, as well as the molecular orbital structure. The quality of the retrieval of the molecular features relies, therefore, on the fairness of the SFA in describing the HHG process. The first step of HHG is the tunnel ionization, which constitutes a fundamental process in strong-field interactions and has been extensively studied in atoms [16, 32–34] and extended to molecules [35–39] in the SFA. Notably, the SFA

neglects the shape of the molecular potential barrier at ionization, thus, describing the ionization from each molecular site on equal foot. However, for the actual potential shapes, the elimination of the ion potential barriers inside the molecule leads to extended molecular orbitals, where part of the electron wavefunction is effectively separated from the edge of the tunnel ionization barrier. Recently, Liu and Liu have suggested treating the ionization from each atomic site differently, though by considering different effective barrier widths [39]. Finally, Labeye *et al* have studied the implications of the internal molecular barrier in the semiclassical description of HHG from molecules [40].

In this work, we demonstrate that the standard tunnel ionization picture has to be revised in the molecular case. As we will show, the tunnel ionization probability is affected by the separation of part of the electronic wavefunction from the potential barrier, introducing a molecular-site specificity in the electron ionization. As tunnel ionization is a key ingredient in HHG, the signature of its site-specificity is evidenced in the details of the high-order harmonic and photoelectron spectra. The total ionization rates, as given by the molecular Ammosov–Delone–Krainov theory [35] and the weak-field asymptotic theory [37], are little affected. As a practical consequence, the orbital information encoded in the harmonic spectrum is substantially distorted from the prediction of the standard SFA models. We, however, show that the proper description of electron tunneling from molecular orbitals can be incorporated into these existing models with the introduction of a modified molecular form factor. The consequences of this study are two-fold: on the one side, it modifies the interpretation of the harmonic signal in HHS, since the retrieved orbital is not a raw image of the actual molecular wavefunction; on the other side, it demonstrates that the HHG spectrum shows well-resolved signatures of the tunnel-ionization site-specificity. Finally, our considerations lead to a derivation of a molecular SFA HHG model with quantitative accuracy.

The article is organized as follows. First, we describe our theoretical approach to the site-dependent tunneling ionization and its incorporation in the strong field approximation formalism. Second, we present the results for HHG and photoelectron spectra obtained from our model compared to those from the exact TDSE calculations, and finally we conclude.

## 2. Theoretical approach to site-dependent tunnel ionization in molecules

We consider as the subject of study the hydrogen molecular ion ( $\text{H}_2^+$ ) in a linearly polarized laser field, since the exact TDSE calculations are affordable for this simple case [41]. The molecule is assumed to lie parallel to the polarization direction, a feasible scenario after using laser alignment techniques [42]. We also neglect nuclei dynamics, which is reasonable for time scales of few femtoseconds [43]. We consider an 800 nm wavelength laser pulse described as  $E(t) = E_0 \sin^2(\pi t/\tau) \sin(\omega_0 t)$ ,  $\omega_0$  being the laser carrier frequency,  $\tau = 10.67$  fs, which leads to 3.88 fs full-width at half maximum (FWHM) in intensity, and a peak intensity of  $3.5 \times 10^{14}$  W cm<sup>-2</sup>. The Hamiltonian governing the interaction is

$$H(t, \mathbf{r}) = \frac{\mathbf{p}^2(t)}{2m} + V_M(\mathbf{r}), \quad (1)$$

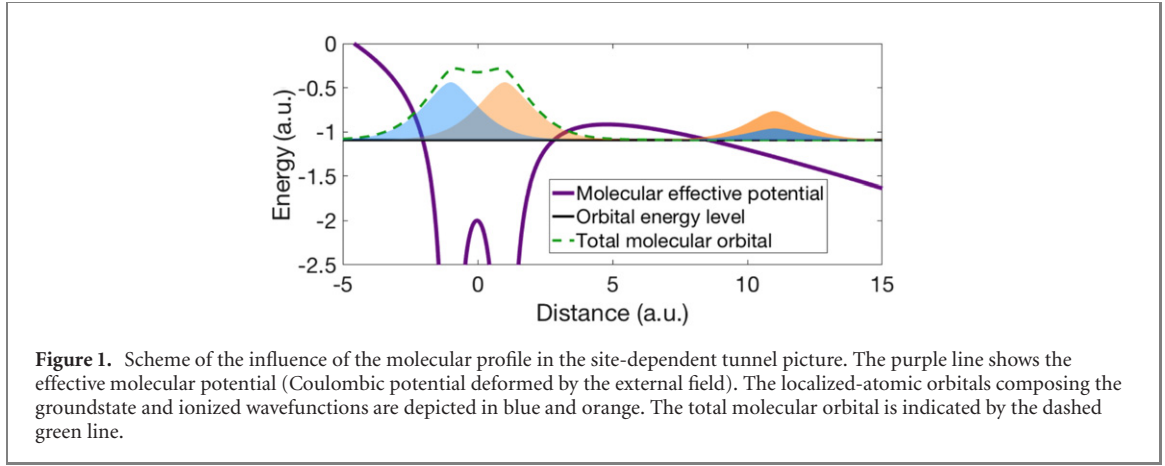
where  $m$  is the electron mass, and

$$V_M(\mathbf{r}) = -\frac{q^2}{\sqrt{\rho^2 + (z + R/2)^2}} - \frac{q^2}{\sqrt{\rho^2 + (z - R/2)^2}} \quad (2)$$

is the Coulomb potential in cylindrical coordinates ( $q$  is the electron charge, and  $R$  is the internuclear distance).  $\mathbf{p}$  is the kinetic momentum operator,  $\mathbf{p} = -i\hbar\nabla - (q/c)A(t)\mathbf{e}_z$ ,  $A(t)$  being the vector potential of the laser field, polarized in the  $z$  direction. As shown in appendix A, the ground-state of  $\text{H}_2^+$  is a symmetric orbital that can be described in terms of atomic orbitals centered at the molecular ion sites, as

$$\phi_0(\mathbf{r}) = C \left[ \chi_0 \left( \mathbf{r} - \frac{R}{2} \mathbf{e}_z \right) + \chi_0 \left( \mathbf{r} + \frac{R}{2} \mathbf{e}_z \right) \right], \quad (3)$$

where  $\chi_0(\mathbf{r})$  is a localized wavepacket, described as a linear combination of Gaussian orbitals [44, 45], and  $C$  is the normalization factor. The harmonic spectrum is calculated from the Fourier transform of the mean dipole acceleration along the  $z$  axis,  $a_z(t) = \langle \hat{a}_z \rangle = \langle -(1/m)\partial V_M/\partial z \rangle$ . The TDSE is integrated using the Crank–Nicholson algorithm in the finite differences scheme.



Molecular HHG can be described using an extension of the atomic SFA models. We use the SFA + approach, which provides a quantitative accurate reproduction of the harmonic spectra in atoms [21]. The extension to molecules leads to the following expression for the dipole acceleration (see appendix A):

$$a_z(t) = -\frac{i}{\hbar} \int_{\mathbf{P}} \alpha_+(\mathbf{P}) \langle \chi_0 | \hat{a}_z | \mathbf{P} \rangle | \mathcal{F}_s[\mathbf{p}'(t)] |^2 \times \int_{t_0}^t e^{i\frac{1}{\hbar} S(\mathbf{P}, t, t_1)} e^{-i\frac{\epsilon_0}{\hbar}(t_1 - t)} V_F(\mathbf{P}, t_1) \times \eta(\mathbf{P}, t_1) d t_1 d \mathbf{P}, \quad (4)$$

where  $\epsilon_0$  is the bound state energy,  $\eta(\mathbf{P}, t_1)$  is the transition matrix element (see equation (A15) in appendix A) and  $\mathcal{F}_s[\mathbf{p}'(t)] = C \left[ e^{i\frac{1}{\hbar} p'_z(t) \frac{R}{2}} + e^{-i\frac{1}{\hbar} p'_z(t) \frac{R}{2}} \right]$  is the molecular form factor, arising from the application of the translation operator to the different molecular ion sites [46].  $\mathbf{P}$  is the electron's canonical momentum,  $S(\mathbf{P}, t, t_1)$  is the ionized electron's action,  $|\mathbf{P}\rangle$  are free-electron wavefunctions [47] and  $V_F(\mathbf{P}, t) = -(q/mc)\mathbf{A}(t) \cdot \mathbf{P} + (q^2/2mc^2)A^2(t)$ . We shall refer to equation (4) using the form factor  $\mathcal{F}_s[\mathbf{p}'(t)]$  as the standard SFA (s-SFA). According to reference [4],  $\mathbf{p}'$  is interpreted as the kinetic momentum of the recolliding electron, as seen from the bottom of the molecular potential well, i.e.  $p'_z(t) = \text{sign}\{p_z(t)\} \sqrt{p_z^2(t) + 2mI_p}$ ,  $p_z(t)$  being the kinetic momentum of the recolliding electron outside the molecular well, and  $I_p$  being the bound-state ionization energy. This description has been validated by exact calculations using Coulomb–Volkov wavefunctions [48]. Finally, the term  $\alpha_+(\mathbf{P}) = (P^2/2m - \epsilon_0)/\Delta_s$  in equation (4) is a prefactor necessary to describe the harmonic yield with quantitative accuracy [21], being  $\Delta_s$  the ground-level Stark shift at the instant of recollision. As it is well-known, the form factor  $\mathcal{F}_s[\mathbf{p}'(t)]$  in equation (4) is responsible for the interference pattern in the harmonic spectrum, dubbed as *structural minimum* [4].

## 2.1. Introduction of the site-depending tunneling

As pointed out above, a main assumption in SFA is to neglect the form of the Coulomb potential upon ionization and, therefore, it oversimplifies the nature of the tunnel barrier [33]. In figure 1 the molecular potential is depicted showing that the inner potential barrier is nonexistent. As a consequence, one of the wavepackets  $\chi_0$  in equation (3) is effectively separated from the outer tunneling barrier, reducing its ionization probability. According to the derivation in appendix B, the SFA description of molecular ionization should be modified to include this site-dependent tunnel probability. To incorporate it, we reformulate the decomposition of the molecular orbital, equation (3), as

$$\phi_0(\mathbf{r}) = C \left[ T \left( \frac{R}{2}, \pm \frac{R}{2} \right) \chi_0 \left( \mathbf{r} - \frac{R}{2} \mathbf{e}_z \right) + T \left( -\frac{R}{2}, \pm \frac{R}{2} \right) \chi_0 \left( \mathbf{r} + \frac{R}{2} \mathbf{e}_z \right) \right], \quad (5)$$

where  $T(z_c, z_b)$  is the ratio between the tunnel amplitude probability of the wavepacket placed at the  $z_c$  molecular site,  $P(z_c, z_b)^{1/2}$ , to the amplitude probability near the  $z_b$  edge of the potential barrier,  $P(z_b, z_b)^{1/2}$ . The positive (negative) sign of  $z_b$  corresponds to ionization through a barrier located at the right (left) side of the molecule. According to appendix B,

$$T(z_c, z_b) = \left[ \frac{P(z_c, z_b)}{P(z_b, z_b)} \right]^{1/2} = \frac{|\chi_0(z_b)|}{|\chi_0(z_c)|}. \quad (6)$$

$T(z_c, z_b)$  in equation (5) modulates the SFA ionization, so that the wavepackets that are located further away from the barrier have less probability to be ionized, as depicted schematically in figure 1.

Thus, due to this different probability of tunneling, the wavepackets ionized from each molecular site have a different amplitude. The site-specific tunnel probability,  $T(z_c, z_b)$ , can be included in equation (4) redefining the s-SFA form factor,  $\mathcal{F}_s$ , as a new site-dependent tunneling (SDT) version. This new form factor not only accounts for the application of the translation operator but also for the consequences of site-dependent tunneling in the ionized wavepacket:

$$\mathcal{F}_{\text{SDT}}^{\pm}[p'_z(t)] = C \left[ T\left(-\frac{R}{2}, \pm\frac{R}{2}\right) e^{\frac{i}{\hbar} p'_z(t) \frac{R}{2}} + T\left(\frac{R}{2}, \pm\frac{R}{2}\right) e^{-\frac{i}{\hbar} p'_z(t) \frac{R}{2}} \right]. \quad (7)$$

Note, however, that the modulus of the  $\mathcal{F}_{\text{SDT}}^{\pm}$  form factor is the same regardless the position  $\pm R/2$  of the barrier. We can, therefore, drop the superindices ( $\pm$ ) when substituting  $\mathcal{F}_s$  by  $\mathcal{F}_{\text{SDT}}$  in equation (4). After this substitution, we shall refer to equation (4) as the site-dependent tunneling SFA (SDT-SFA).

The interpretation of the site-dependening tunneling is probabilistic, as given in appendix B, right after equation (B6). Basically, equation (B5) is the compound probability of two independent events: (i) the particle being located at the edge of the barrier and (ii) the tunnel ionization of a particle in contact with the internal edge of the barrier. The final probability is the product of the probabilities these two events. The compound event appears intrinsically in the TDSE calculations, but it is not well described within the strong field approximation models, as they neglect the shape of the potential, and, therefore the information of the location of the inner edge of the tunneling barrier.

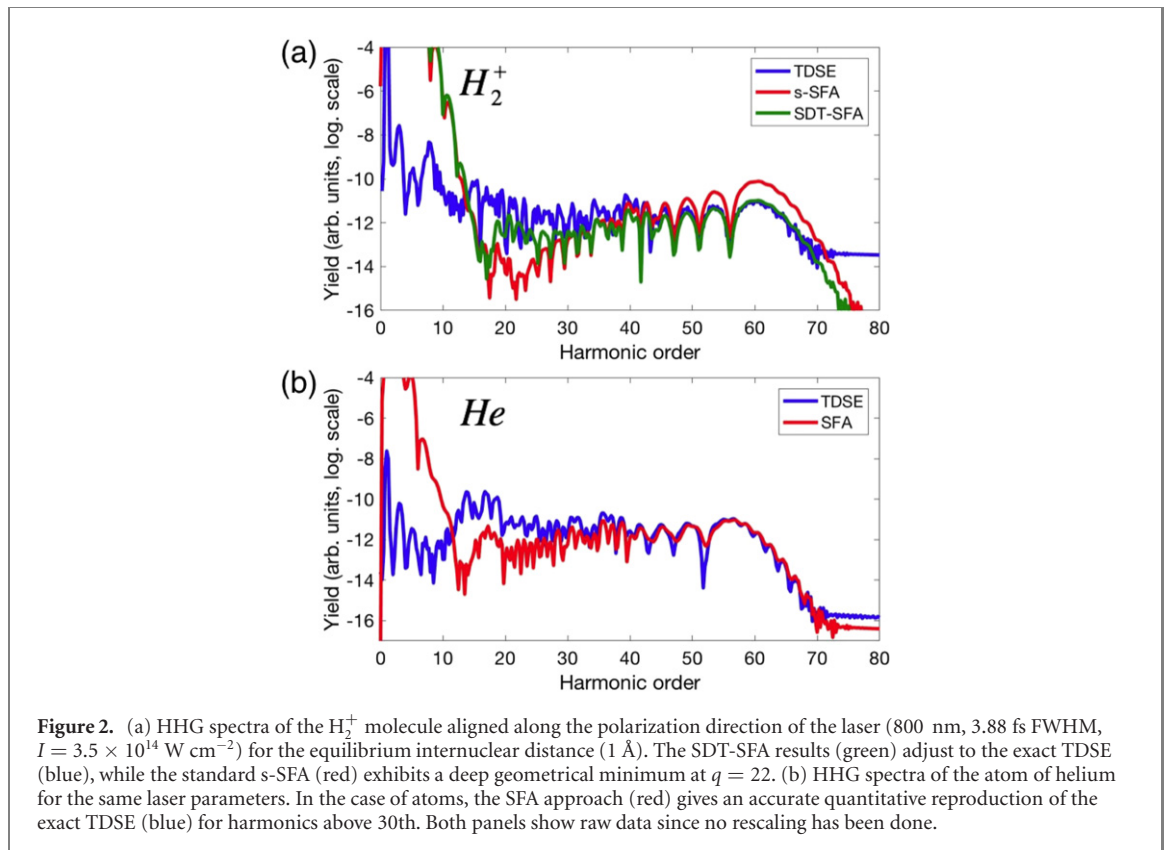
### 3. Signatures of the site-dependent tunnel-ionization in $\text{H}_2^+$

In this section, we show the comparison between the HHG spectra obtained from the TDSE, the standard SFA and the site-dependent tunneling SFA using the form factor  $F_{\text{SDT}}$  in equation (7). Second, we show the comparison of the HHG spectra calculated from the s-SFA, the SDT-SFA and the TDSE, for different laser parameters and different molecular internuclear distances, showing a much better agreement of SDT-SFA with the TDSE. Next, we extract  $F_{\text{SDT}}$  form factor directly from the TDSE calculation, and compare it with equation (7), showing that the TDSE results are consistent with our interpretation. Afterwards, we present results where macroscopic propagation in a molecular target is taken into account, demonstrating that the spectral signature of the site-dependent tunneling is resilient to propagation. Finally, we show the suitability of the SDT-SFA in the photoelectron spectrum emitted in  $\text{H}_2^+$ .

#### 3.1. High harmonic spectra: single-molecule response

Figure 2(a) shows the quantitative comparison between the HHG spectrum obtained with the exact TDSE (blue), the s-SFA (red), and the SDT-SFA (green) as raw data (no rescaling is done). The spectra show the typical feature of non-perturbative harmonic generation: a *plateau* followed by a cut-off. The s-SFA shows a qualitative recovery of the harmonics close to the cut-off (though, one order of magnitude quantitative error), but a serious qualitative and quantitative departure (up to four orders of magnitude) in the *plateau*, most evident for harmonic orders below the 45th. These trends are also present in calculations for different laser parameters, as it will be shown below. In contrast, this clear departure is not found in HHG from atoms. To illustrate this, we show in figure 2(b) the same comparison as figure 2(a) but for the helium atom, with an ionization potential similar to  $\text{H}_2^+$ . The helium case corresponds to equation (4), replacing  $\mathcal{F}_s[\mathbf{p}'(t)]$  by the atomic form factor, which is equal to one, and substituting the matrix elements and orbital energies accordingly. Figure 2(b) shows that the agreement between atomic SFA and TDSE is excellent even for harmonics well into the cutoff (in this case, for harmonic orders above the 30th). The departure of the SFA description for the lowest frequencies (below 30th harmonic in this case) is a known artifact of the strong field approximation. Thus, while the SFA offers an accurate description of HHG in atoms, it substantially fails to describe HHG in  $\text{H}_2^+$ . Such departure reveals that some relevant information is missing in the s-SFA formulation, equation (4).

The presence of the sharp minimum in the s-SFA spectra has been reported before [49, 50], and it corresponds to the molecular structural minimum mentioned above. For  $\text{H}_2^+$  at equilibrium internuclear distance, interacting with an 800 nm-wavelength laser pulse, the structural minimum is centered at the 22nd harmonic, extending up to the 49th harmonic (as found when imposing the condition  $\lambda_{\text{dB}} = 4R/3$ , where  $\lambda_{\text{dB}}$  is the deBroglie wavelength of the recombining electron). However, in contrast to the s-SFA results, the exact integration of the TDSE shown in figure 2(a) shows a weak trace of this interference. This attenuation of the structural minimum in  $\text{H}_2^+$  at the equilibrium internuclear distance, parallel to the laser polarization, has been also evidenced in previous works [4, 6, 50]. On the other hand, it is also known that



the minimum shows up for tilted molecular orientations [6, 25], where the projection of the molecular axis onto the field polarization results into an effective internuclear distance smaller than the equilibrium one. We will later show results of the TDSE for such case, where the structural minimum is clearly observed.

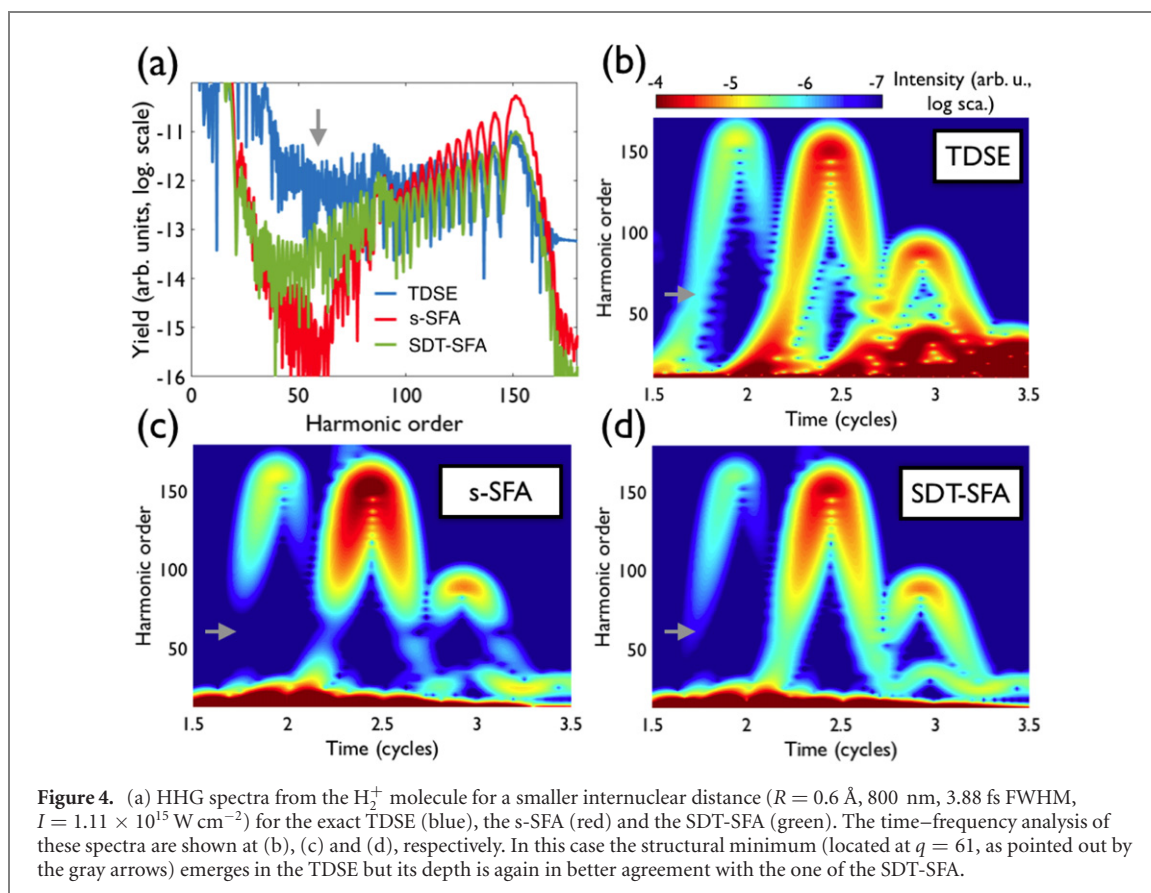
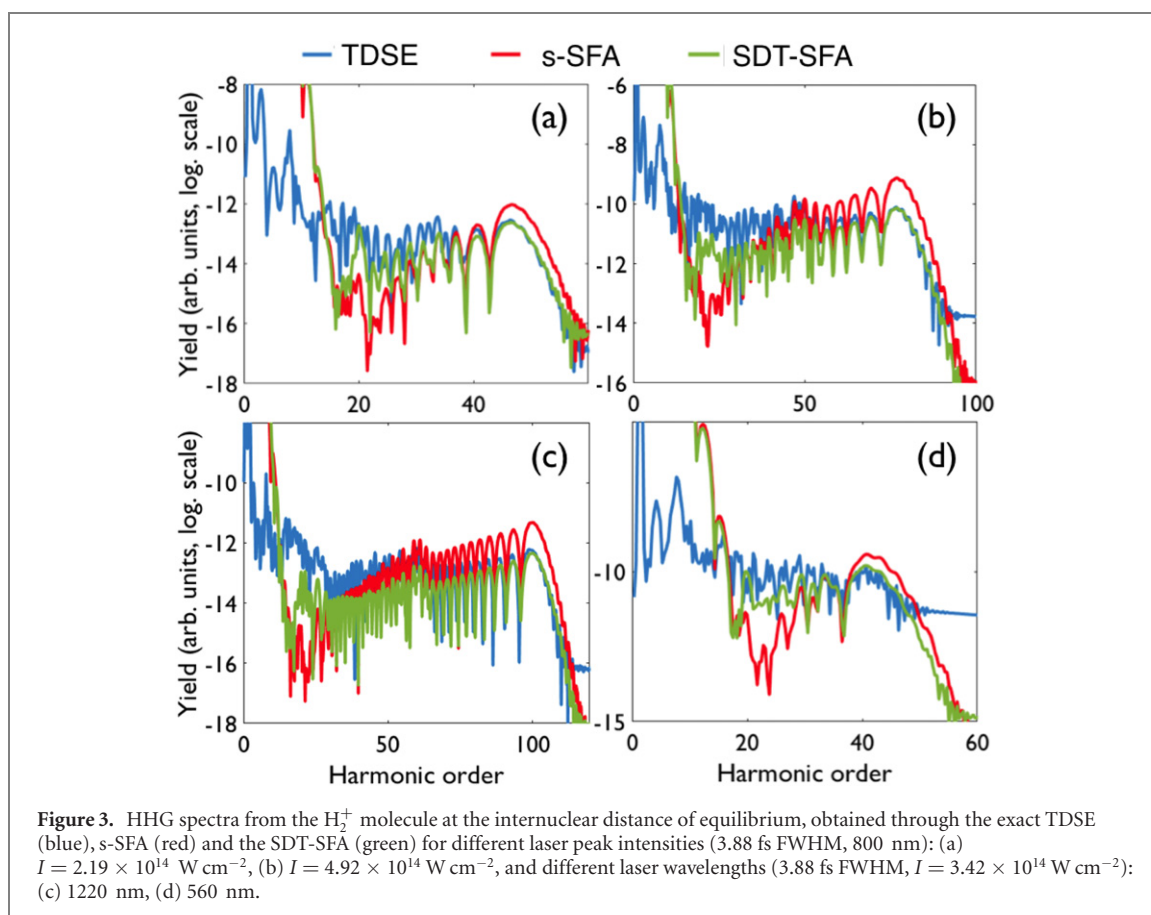
The comparisons of the results of SDT-SFA with s-SFA and the exact TDSE are shown in figure 2(a). In agreement with the TDSE, the SDT-SFA spectrum shows a weaker signature of the structural minimum. Remarkably, also the overall correspondence of SDT-SFA with the TDSE—both quantitative and qualitative—is substantially improved in comparison with the s-SFA. It should be reminded that the departure at the lowest part of the spectral plateau (harmonic orders <30th) is also found in the atomic case (figure 2(b)). As discussed before, it reflects a fundamental inaccuracy of the SFA approach, not connected to the atomic or molecular nature of the species.

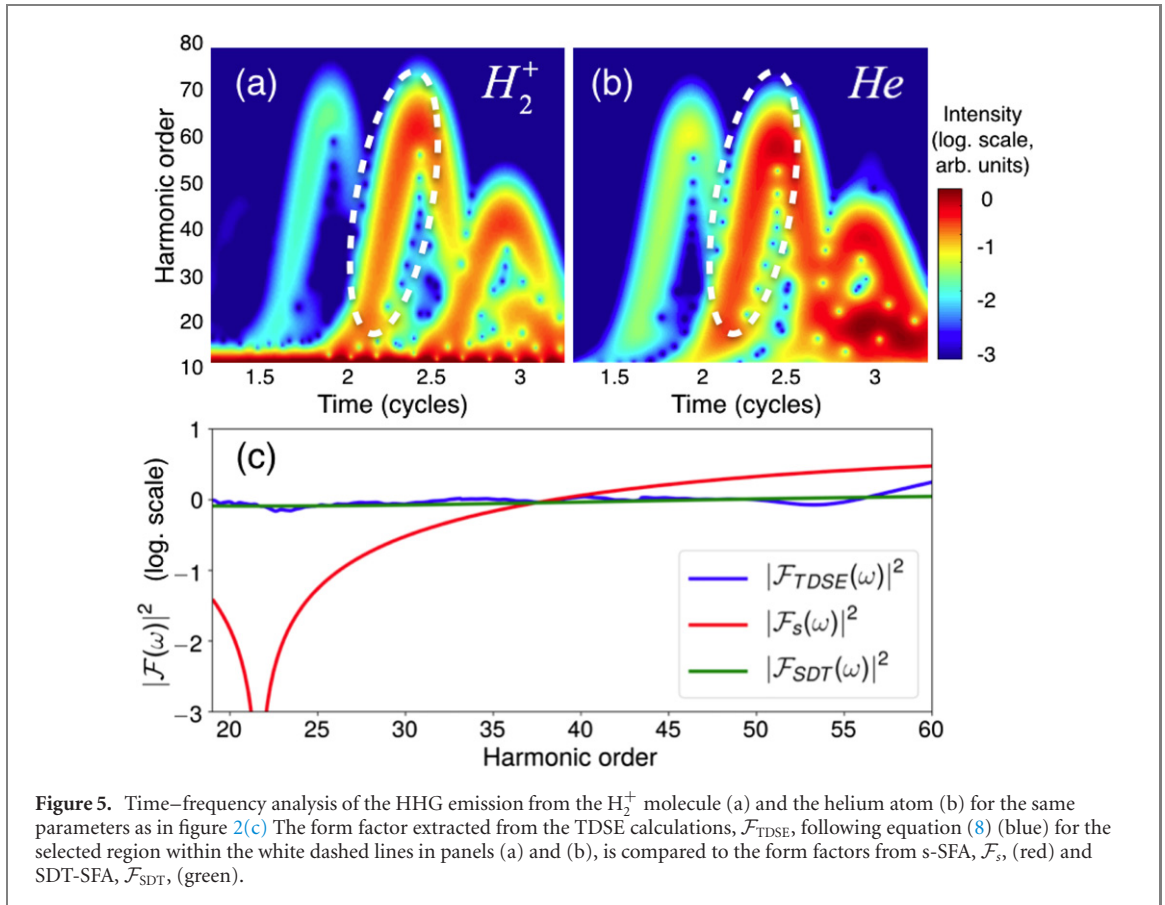
We validate the SDT-SFA model for different laser parameters and internuclear distances. First, in figure 3 we present the HHG spectra considering different laser peak intensities and wavelengths. As a general conclusion, the SDT-SFA (green lines) reproduces very satisfactorily the exact TDSE results (blue lines).

On the other hand, tilted molecules present an effective internuclear distance smaller than the equilibrium one:  $R_{\text{eff}} = R_{\text{eq}} \cos \theta$ . Under such configurations the ratio of the tunnel probability of a distant wavepacket (equation (6)) increases, and the structural minimum shows up. Thus, we have chosen a smaller internuclear distance in order to study the capability of the SDT-SFA to reproduce the well-known structural minimum [4] in a configuration where it is evidenced. In figure 4 we present the HHG spectra and its corresponding time–frequency analysis for a different internuclear distance,  $R = 0.6 \text{ \AA}$ , which would correspond to a rotation of  $\theta = 53^\circ$ . The minimum appears in the TDSE spectrum at the harmonic order  $q = 61$ , approximately. While the standard SFA fails in predicting the minimum’s depth, the SDT-SFA provides a better approach. Note that the time–frequency description of the SDT-SFA has also a better coincidence with the TDSE results. As pointed out before, the departures for harmonic orders below 30th are a consequence of the SFA, also present in the atomic case (figure 2(b)).

### 3.2. Extraction of the form factor from the harmonic spectra

A strong evidence of site-specificity in molecular tunnel ionization can be found directly from the exact TDSE using high-harmonic spectroscopy techniques to retrieve the form factor from the TDSE and compare it with our SDT proposal, equation (7). For that purpose, we follow the same philosophy used in tomographic studies and we compare the HHG molecular emission to that from an atom with similar





ionization potential. Figures 5(a) and (b) show the time–frequency analysis corresponding to the TDSE harmonic spectra of figures 2(a) and (b), for  $H_2^+$  and He respectively. These maps reveal that each harmonic in the *plateau* is emitted in a discrete series of bursts, corresponding to the rescattering of the different electron trajectories, the well-known short and long trajectories [19, 51]. We have selected the harmonic emission corresponding to a single rescattering event emitted from the short trajectories during a half-cycle of the driving field (as indicated with white dashed circles in figures 5(a) and (b)) and calculated the associated spectral amplitudes for both species:  $a_{z,\text{He}}(\omega)$  and  $a_{z,H_2^+}(\omega)$ . Then, the form factor can be retrieved as

$$|\mathcal{F}_{\text{TDSE}}(\omega)|^2 \simeq a_{z,H_2^+}(\omega)/a_{z,\text{He}}(\omega). \quad (8)$$

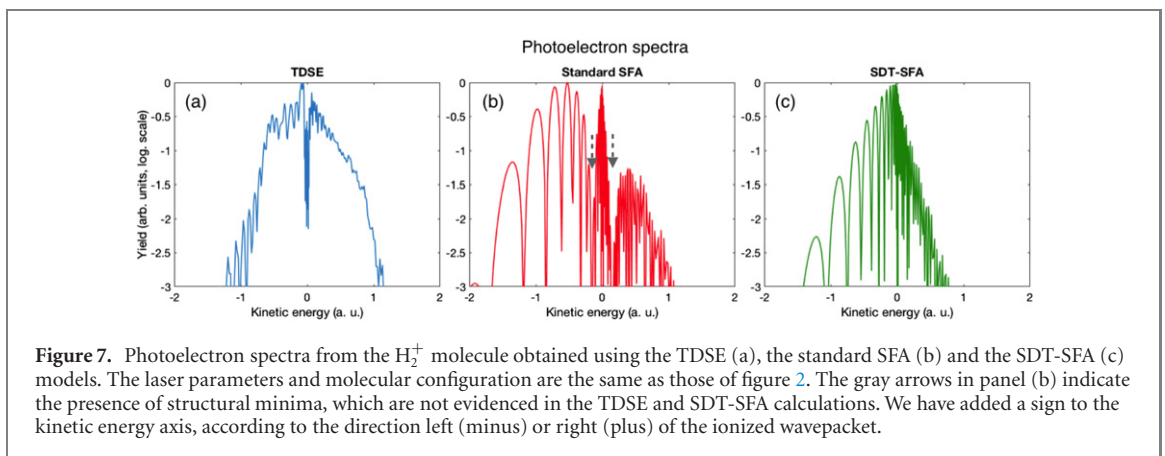
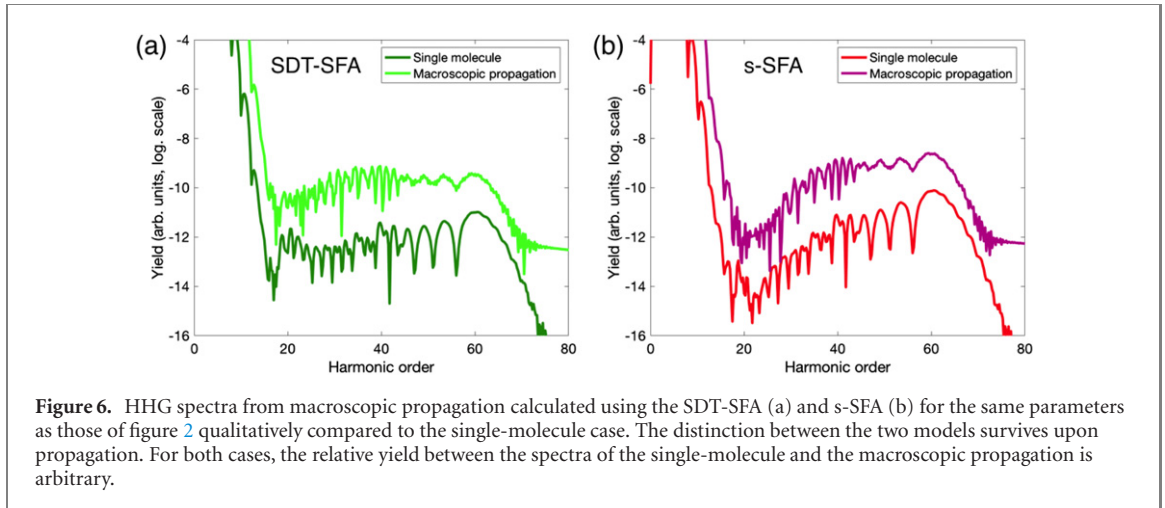
This assumption follows from the SFA saddle point method, which applied to the integrals in equation (4) leads to a simple identification of the harmonic spectral amplitudes and the quasiclassical electron trajectories [19]. After the saddle-point analysis, the Fourier transform of equation (4) can be cast into a simple expression [52]:

$$a_z(\omega) \propto \sum_{st} |\mathcal{F}_{\text{TDSE}}(p'_{st})|^2 \xi(\mathbf{P}_{st}, t_{st}, t_{1,st}) \eta(\mathbf{P}_{st}), \quad (9)$$

where  $\eta$  includes the terms containing the electronic wavefunctions,  $\chi_0$ , and  $\xi$  includes the remaining terms, except the form factor  $\mathcal{F}_{\text{TDSE}}$ , which is to be determined from the TDSE. The summation in equation (9) extends over all saddle points,  $st$ , each representing a recolliding electron trajectory responsible of the harmonic emission at frequency  $\omega = p_{st}^2/2m\hbar$ . Each rescattering corresponds to a term in the summation in equation (9), therefore, in selecting a single rescattering event from the TDSE time–frequency maps, we are isolating a single term in the summation. Thus, in this case, equation (8) follows directly from equation (9).

Figure 5(c) shows the comparison of the form factor extracted from the TDSE,  $\mathcal{F}_{\text{TDSE}}$ , using equation (8), against the s-SFA form factor,  $\mathcal{F}_s$ , and the modified SDT-SFA version,  $\mathcal{F}_{\text{SDT}}$ . The excellent agreement between  $\mathcal{F}_{\text{SDT}}$  and  $\mathcal{F}_{\text{TDSE}}$  strongly supports equation (7) and, therefore, our interpretation of distant tunneling as well as molecular site-dependent ionization. We note that evidences of the site-dependent ionization due to the distant tunneling may also be found implicitly in photoionization studies where electron localization leads to an asymmetric molecular dissociation [53–55].





### 3.3. Signature of the site-dependent tunnel-ionization in the propagated high-harmonic signal

We introduce the macroscopic picture in order to gain insight about the survival of the features of the single-molecule spectrum in an experimental situation. It is known that macroscopic phase-matching can strongly influence the HHG spectrum measured in an experiment [56, 57]. In previous works, the experimental HHG spectrum has been used to retrieve the molecular orbital using tomographic techniques [13]. Therefore, it is important to know if the SDT features of the single-molecule HHG spectrum are also found in the macroscopic picture.

The macroscopic simulation of HHG in  $H_2^+$  is based on the electromagnetic field propagator [58], in which we discretize the target (molecular gas jet) into elementary radiators. The dipole acceleration of each elementary source is computed using the SFA models described by equations (A21) and (B10). We assume that the harmonic radiation propagates with the vacuum phase velocity, which is a reasonable assumption for high-order harmonics. The low-density molecular gas jet, flowing along the perpendicular direction to the beam propagation is modelled as a Gaussian distribution of  $200 \mu\text{m}$  at full width half maximum, and with a peak pressure of 5 torr. The driver field has a Gaussian profile with a beam waist at the focus position of  $120 \mu\text{m}$ . All molecules are assumed to be oriented parallel to the polarization direction.

In figure 6, we show a qualitative comparison between the single-molecule and macroscopic HHG spectra for the SDT-SFA and the s-SFA models. The main features of the single-molecule spectrum calculated using both models are conserved in the macroscopic case and, although the shape of SDT-SFA spectra suffers modifications upon macroscopic propagation, it remains clearly distinct from the s-SFA results. Our results show that the main properties of the single-molecule HHG spectrum shown in figure 2 survive after macroscopic propagation.

### 3.4. Signatures of the site-dependent tunnel-ionization in the photoelectron spectra

In this section, we calculate the photoelectron spectra from the  $H_2^+$  molecule along the  $z$  coordinate,  $|\delta\Psi(P_z)|^2$ , after the interaction with the pulse. In figure 7, we present the photoelectron spectra calculated from the TDSE (a), standard SFA (b) and SDT-SFA (c) models. We show that, while the SDT-SFA spectrum

exhibits a better agreement with the TDSE spectrum, the standard SFA spectrum presents two deep minima [59]. The origin of those minima is the interference of the ionization from each atomic orbital. The SDT-SFA, on the other hand, includes the side-dependency of the ionization, which sharply decreases the structural interference.

Note that, in order to obtain the photoelectron spectrum from the TDSE, we have filtered the fundamental state from the final wavefunction. As a result, an artificial minimum at the lowest kinetic energies appears in the spectrum.

#### 4. Conclusions

In conclusion, we have demonstrated that the standard picture of tunnel ionization needs to be modified for the non-atomic case. Our exact computations of the TDSE in  $H_2^+$  reveal that wavepacket portions located at the ion sites separated from the barrier ionize with lower probability. We propose a corrected molecular form factor to implement into the existing strong-field models. The new form factor agrees extremely well with the one extracted from the exact TDSE solution, and it improves both the HHG and photoelectron spectra. In addition, the signatures of the site-dependent tunneling are present in the HHG spectra for different laser parameters and molecular internuclear distances. We show that those spectral signatures are also present when the macroscopic phase-matching is taken into account. We believe that the implementation of the site-dependent corrections in the retrieval algorithms will improve substantially the accuracy of high-harmonic spectroscopy measurements, as well as tomographic orbital image reconstruction.

#### Acknowledgments

C H-G and L P acknowledge support from Junta de Castilla y León and European Union, FEDER (Project SA287P18), Spanish MICINN (FIS2016-75652 P). C H-G acknowledges Ministerio de Ciencia, Innovación y Universidades for a Ramón y Cajal contract (RYC-2017-22745), co-funded by the European Social Fund. L R acknowledges support from Ministerio de Educación, Cultura y Deporte (FPU16/02591). A P acknowledges support from Comunidad de Madrid through the TALENTO program with Reference No. 2017-T1/IND-5432. We thankfully acknowledge the computer resources at MareNostrum and the technical support provided by Barcelona Supercomputing Center (FI-2019-1-0013).

#### Appendix A. Standard SFA description of molecular HHG

In this appendix, we present the standard (s-SFA) description of molecular high-order harmonic generation (HHG) within the strong-field approximation (SFA).

The interaction of a one-electron system with electromagnetic radiation, in the dipole approximation, is governed by the Hamiltonian

$$H(t) = \frac{p^2}{2m} + V_C(\mathbf{r}) + V_F(\mathbf{P}, t), \quad (\text{A1})$$

where  $V_C(\mathbf{r})$  is the Coulombic potential and  $V_F(\mathbf{P}, t) = -(q/mc)\mathbf{A}(t) \cdot \mathbf{P} + (q^2/2mc^2)A^2(t)$ , being  $\mathbf{A}(t)$  the vector potential of the laser field,  $\mathbf{P}$  the canonical momentum and  $m$  and  $q$  the mass and the charge of the electron, respectively.

The standard SFA approach is based in the assumption that, once ionized, the electron dynamics is governed by the field interaction, neglecting the effect of the ion Coulombic potential. The SFA propagator is written as [18]

$$G_{\text{SFA}}(t, t_0) = G_0(t, t_0) + \frac{1}{\hbar} \int_{t_0}^t G_F(t, t_1) V_F(t_1) G_0(t_1, t_0) dt_1, \quad (\text{A2})$$

where  $G_0$  is the propagator of the non-interacting Hamiltonian,  $H_0 = \frac{p^2}{2m} + V_C(\mathbf{r})$ , and  $G_F$  is the propagator of the free electron in the electromagnetic field, described by  $H_F(t) = \frac{p^2}{2m} + V_F(t)$ . Within SFA, the electron's wavefunction is given by  $|\psi(t)\rangle = iG_{\text{SFA}}(t, t_0)|\phi_0\rangle$ , with  $|\phi_0\rangle$  an initial bound-state of the system. We can split the wavefunction into two terms,  $|\psi(t)\rangle = |\phi_0(t)\rangle + |\delta\psi(t)\rangle$ , with  $|\phi_0(t)\rangle = iG_0(t, t_0)|\phi_0\rangle$  the bound electron evolving in the absence of the field, and

$$|\delta\psi(t)\rangle = \frac{i}{\hbar} \int_{t_0}^t G_F(t, t_1) V_F(t_1) G_0(t_1, t_0) |\phi_0\rangle dt_1 \quad (\text{A3})$$

the electron in the continuum. Defining  $\epsilon_0$  as the bound-state energy, we have  $G_0(t_1, t_0)|\phi_0\rangle = -ie^{-i\epsilon_0(t_1-t_0)/\hbar}|\phi_0\rangle$ . For the electron in the continuum, we resort to the Volkov basis [47]:  $|\mathbf{P}(t)\rangle = e^{i\frac{1}{\hbar}S(\mathbf{P}, t, t_0)}|\mathbf{P}\rangle$ , with  $S(\mathbf{P}, t, t_0)$  is the action defined as  $S(\mathbf{P}, t, t_0) = -\frac{1}{2m}\int_{t_0}^t p^2(\tau)d\tau$ , with  $\mathbf{p}(t) = \mathbf{P} - (q/c)\mathbf{A}(t)$ , the kinetic momentum. Accordingly, the free electron propagator can be expressed in the Volkov basis as  $G_F(t, t_1) = -i\int e^{i\frac{1}{\hbar}S(\mathbf{P}, t, t_1)}|\mathbf{P}\rangle\langle\mathbf{P}|d\mathbf{P}$ . Using these definitions in (A3) the ionized electron wavefunction in momentum space reads as

$$\delta\psi(\mathbf{P}, t) = \langle\mathbf{P}|\delta\psi(t)\rangle = -\frac{i}{\hbar}\int_{t_0}^t e^{i\frac{1}{\hbar}S(\mathbf{P}, t, t_1)}e^{-i\epsilon_0(t_1-t_0)/\hbar}V_F(\mathbf{P}, t_1)\eta(\mathbf{P})d t_1, \quad (\text{A4})$$

where  $\eta(\mathbf{P})$  is the transition matrix element

$$\eta(\mathbf{P}) = \langle\mathbf{P}|\frac{C_F}{r^n}|\phi_0\rangle, \quad (\text{A5})$$

where the factor  $C_F/r^n$  is a Coulomb correction [33] that improves the quantitative accuracy of the SFA description [21, 60], with  $C_F = [4|\epsilon_0|/(|q|E_0)]^2$  and  $n = (Zq^2/\hbar)\sqrt{m/2|\epsilon_0|}$ ,  $Z$  being the charge of the atomic or ionic core ( $Z = 1$  for the hydrogen atom and  $Z = 2$  for  $\text{H}_2^+$ ).

The coherent radiation spectrum is proportional to the Fourier transform of the mean acceleration. Since we are interested only in high-harmonics, we compute the complex acceleration amplitude as [21]

$$\langle\mathbf{a}(t)\rangle = \int \frac{p^2/2m - \epsilon_0}{\Delta_s}\mathbf{a}(\mathbf{P}, t)d\mathbf{P}, \quad (\text{A6})$$

with

$$\mathbf{a}(\mathbf{P}, t) = (1/m)\langle\phi_0(t)|(-\nabla V_C)|\mathbf{P}\rangle\delta\psi(\mathbf{P}, t) \quad (\text{A7})$$

being the contribution of each Volkov wave to the total acceleration. The prefactor of the integrand in (A6) accounts for the boundstate dressing at the instant of recollision, necessary to describe the harmonic yield with quantitative accuracy,  $\epsilon_0$  being the energy of the bound orbital and  $\Delta_s$  the Stark shift at the instant of recollision [21, 61]. This formulation has been previously applied to atoms successfully, and an example is shown in figure 2(b).

Let us now introduce the extension of this SFA description to molecules, named as standard SFA (s-SFA) in this paper. We describe the molecular hydrogen ion ground state as a linear combination of atomic orbitals (LCAO), with two centers at  $\pm R/2$ ,  $R = 2$  a.u being the internuclear distance. The LCAO are described with a 6-311G Pople basis, and determined from a variational calculation. For the case of  $\text{H}_2^+$  we have three basis in each molecular site,  $\mathbf{R}_i$  defined as

$$\begin{aligned} \Phi_{1s}(\mathbf{r}) &= 0.0254938g_s(33.8650000, \mathbf{r}) + 0.1903730g_s(5.0947900, \mathbf{r}) + 0.8521610g_s(1.1587900, \mathbf{r}) \\ \Phi_{2s}(\mathbf{r}) &= 1.0000000g_s(0.3258400, \mathbf{r}) \\ \Phi_{3s}(\mathbf{r}) &= 1.0000000g_s(0.1027410, \mathbf{r}) \end{aligned} \quad (\text{A8})$$

being  $g_s(\alpha, \mathbf{r}) = (2\alpha/\pi)^{3/4}e^{-\alpha r^2}$ . By using Hartree–Fock, the molecular orbital for the ground state is found to be:

$$\phi_0(\mathbf{r}, t=0) = C\left[\chi_0\left(\mathbf{r} - \frac{R}{2}\mathbf{e}_z\right) + \chi_0\left(\mathbf{r} + \frac{R}{2}\mathbf{e}_z\right)\right], \quad (\text{A9})$$

where  $C$  is the normalization factor and

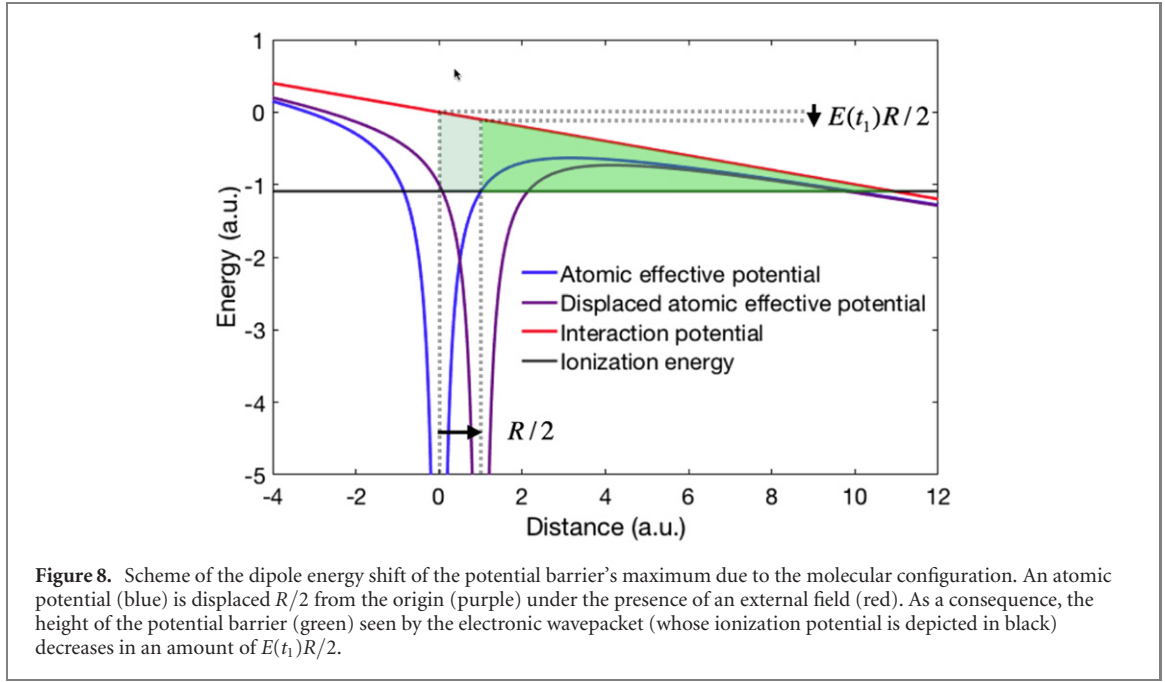
$$\chi_0(\mathbf{r}) = 0.1937\Phi_{1s}(r) + 0.3990\Phi_{2s}(r) + 0.0484\Phi_{3s}(r). \quad (\text{A10})$$

Thus, the time-dependent molecular orbital can be expressed as

$$\phi_0(\mathbf{r}, t) = Ce^{-i\epsilon_0(t-t_0)/\hbar}\left[\chi_0\left(\mathbf{r} - \frac{R}{2}\mathbf{e}_z\right) + \chi_0\left(\mathbf{r} + \frac{R}{2}\mathbf{e}_z\right)\right] \rightarrow |\phi_0(t)\rangle = C\left(e^{-\frac{i}{\hbar}\hat{p}_z(t)\frac{R}{2}} + e^{\frac{i}{\hbar}\hat{p}_z(t)\frac{R}{2}}\right)|\chi_0(t)\rangle, \quad (\text{A11})$$

where  $|\chi_0(t)\rangle = e^{-i\epsilon_0(t-t_0)/\hbar}|\chi_0\rangle$  and the binding energy is calculated to be  $\epsilon_0 = -29.65$  eV ( $-1.09$  a.u.). Consequently, the ionized wavefunction can be written as

$$|\delta\Psi(t)\rangle = C\left(e^{-\frac{i}{\hbar}\hat{p}_z(t)\frac{R}{2}} + e^{\frac{i}{\hbar}\hat{p}_z(t)\frac{R}{2}}\right)|\delta\psi(t)\rangle, \quad (\text{A12})$$



where  $\hat{p}'_z(t)$  is the projection of the kinetic momentum operator, as seen from the molecular well (see below). The expression for  $|\delta\psi(t)\rangle$  in equation (A4) should be corrected to take into account the dipole energy shift of the barrier's maximum, as the molecule is an extended object. For a symmetric aligned molecule of length  $R$ , the change in the barrier height affects the ionization probability by a fraction  $W_{\text{ADK}}[\epsilon(R/2, t_1)]/W_{\text{ADK}}(\epsilon_0)$ , where  $\epsilon(R/2, t_1) = \epsilon_0 - q|E(t_1)|R/2$  is the ionization potential as seen from the top of the barrier at the ionization time  $t_1$  (see figure 8).  $W_{\text{ADK}}$  is the Ammosov–Delone–Krainov the ionization rate [34]

$$W_{\text{ADK}}(\epsilon) = \left(\frac{2e}{n^*}\right)^{2n^*} \frac{1}{2\pi n^*} \frac{(2\ell + 1)(\ell + |m|)!}{2^{|m|}|m|!(\ell - |m|)!} |\epsilon| \left[\frac{3E_0}{\pi(2|\epsilon|)^{3/2}}\right]^{1/2} \times \left[\frac{2}{E_0}(2|\epsilon|)^{3/2}\right]^{2n^* - |m| - 1} e^{-\left(\frac{2(2|\epsilon|)^{3/2}}{3E_0}\right)}, \quad (\text{A13})$$

where  $e$  is the Euler number,  $m$  and  $\ell$  are the quantum numbers of the atomic orbital,  $E_0$  is the amplitude of the electric field and

$$n^*(\epsilon) = \frac{Zq^2}{\hbar} \sqrt{\frac{m}{2|\epsilon|}}. \quad (\text{A14})$$

We implement this correction including the probability ratio into the bound-to-continuum amplitude probability, thus, redefining  $\eta$  in equation (A4) as

$$\eta(\mathbf{P}, t_1) = \left[\frac{W_{\text{ADK}}[\epsilon(R/2, t_1)]}{W_{\text{ADK}}(\epsilon_0)}\right]^{1/2} \langle \mathbf{P} | \frac{C_F}{r^n} | \chi_0 \rangle, \quad (\text{A15})$$

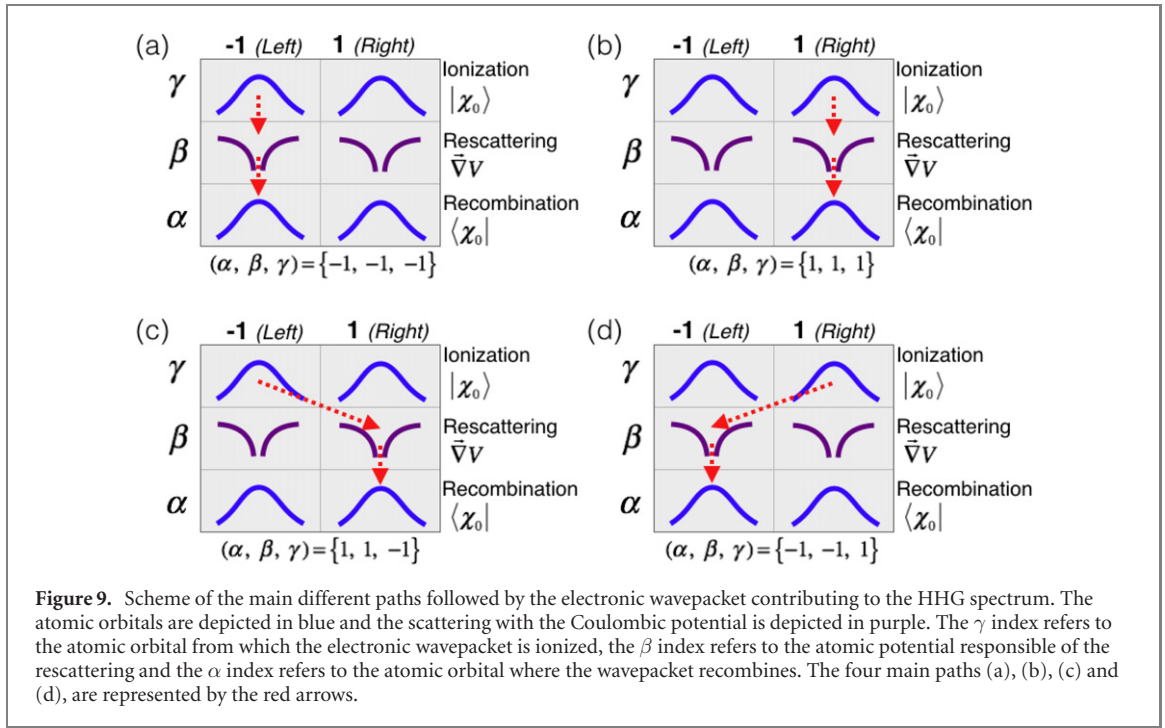
where  $R$  is the equilibrium internuclear distance ( $R = 1.055 \text{ \AA} = 2 \text{ a.u.}$ ).

Equations (A7), (A11) and (A12), lead to the following expression for the dipole acceleration for the  $\text{H}_2^+$  molecule,

$$\mathbf{a}(\mathbf{P}, t) = \frac{C^2}{m} \langle \chi_0(t) | \left( e^{\frac{i}{\hbar} \hat{p}'_z(t) \frac{R}{2}} + e^{-\frac{i}{\hbar} \hat{p}'_z(t) \frac{R}{2}} \right) (-\nabla V_M) | \mathbf{P} \rangle \langle \mathbf{P} | \left( e^{-\frac{i}{\hbar} \hat{p}'_z(t) \frac{R}{2}} + e^{\frac{i}{\hbar} \hat{p}'_z(t) \frac{R}{2}} \right) | \delta\psi(t) \rangle, \quad (\text{A16})$$

$V_M$  where is the Coulomb molecular potential. For  $\text{H}_2^+$ ,  $V_M$  can be written as a superposition of the hydrogen potential at the ionic sites,  $V_M(\mathbf{r}) = V_C(\mathbf{r} - \frac{R}{2}\mathbf{e}_z) + V_C(\mathbf{r} + \frac{R}{2}\mathbf{e}_z)$ . The dipole acceleration  $\mathbf{a}(\mathbf{P}, t)$ , therefore, results from the added contributions of eight different physical paths (see figure 9). Naming  $\{\alpha, \beta, \gamma\}$  the sign of the displacements (+1 for right and -1 for left) of the recombination wavefunction  $\chi_0(\mathbf{r} - \alpha R/2)$ , the rescattering potential  $V_C(\mathbf{r} - \beta R/2)$ , and the ionizing wavefunction  $\chi_0(\mathbf{r} - \gamma R/2)$ , respectively, we rewrite (A16) as the sum over the different paths  $\mathbf{a}(\mathbf{P}, t) = \sum_{\alpha, \beta, \gamma} \mathbf{a}^{\alpha, \beta, \gamma}(\mathbf{P}, t)$ , where

$$\mathbf{a}^{\alpha, \beta, \gamma}(\mathbf{P}, t) = \frac{C^2}{m} \langle \chi_0(t) | \left\{ e^{\frac{i}{\hbar} \hat{p}'_z \alpha \frac{R}{2}} \left[ -\nabla V_C \left( \mathbf{r} - \beta \frac{R}{2} \mathbf{e}_z \right) \right] | \mathbf{P} \rangle \langle \mathbf{P} | e^{-\frac{i}{\hbar} \hat{p}'_z \gamma \frac{R}{2}} \right\} | \delta\psi(t) \rangle. \quad (\text{A17})$$



Our calculations show that the main contributions come from terms with  $\alpha = \beta$ , meaning that the scattering on a potential site is most probably followed by a recombination to the same site's bound wavefunction. Thus, using

$$-\nabla V_C \left( \mathbf{r} - \beta \frac{R}{2} \mathbf{e}_z \right) = e^{-\frac{i}{\hbar} \hat{p}_z \beta \frac{R}{2}} (-\nabla V_C(\mathbf{r})) e^{\frac{i}{\hbar} \hat{p}_z \beta \frac{R}{2}} \quad (\text{A18})$$

we can approximate equation (A17) to

$$\mathbf{a}^{\beta\gamma}(\mathbf{P}, t) \simeq \frac{C^2}{m} \langle \chi_0(t) | [-\nabla V_C(\mathbf{r})] | \mathbf{P} \rangle e^{\frac{i}{\hbar} \hat{p}_z (\beta - \gamma) \frac{R}{2}} \langle \mathbf{P} | \delta \psi(t) \rangle. \quad (\text{A19})$$

For a free electron  $\mathbf{p}'(t) = \mathbf{p}(t) = \mathbf{P} - (q/c)\mathbf{A}(t)$ , with  $\mathbf{A}(t)$  the electromagnetic vector potential. In molecules, however,  $\mathbf{p}'(t)$  describes the kinetic momentum of the free electron at the instant of recombination,  $t$ , as seen from the molecular sites, so it includes the acceleration by the potential well, i.e.  $(1/2m)p'^2(t) = (1/2m)p^2(t) + I_p$ ,  $I_p$  being the molecular ionization potential. Therefore  $p'_z(t) = \text{sign}\{p_z(t)\} \sqrt{p^2(t) + 2mI_p}$  [48]. This correction to the free electron's kinetic momentum is found necessary to recover the correct position of the molecular structural minimum in the harmonic spectrum [4]. Summing over all the relevant paths, the acceleration  $\mathbf{a}(\mathbf{P}, t)$  can be written as

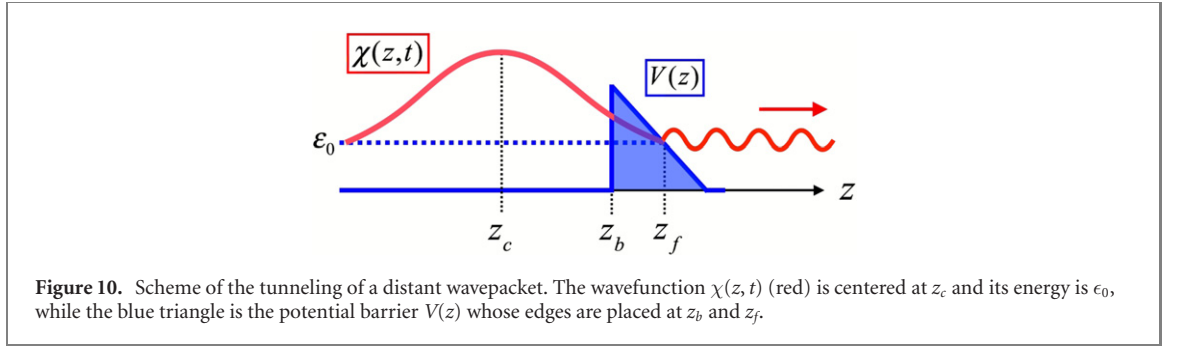
$$\mathbf{a}(\mathbf{P}, t) = \frac{1}{m} |\mathcal{F}_s [p'_z(t)]|^2 \langle \chi_0(t) | (-\nabla V_C) | \mathbf{P} \rangle \delta \psi(\mathbf{P}, t), \quad (\text{A20})$$

where  $\mathcal{F}_s [p'_z(t)] = C(e^{\frac{i}{\hbar} \hat{p}_z(t) \frac{R}{2}} + e^{-\frac{i}{\hbar} \hat{p}_z(t) \frac{R}{2}})$  is the molecular form-factor. Substituting in equation (A6), we finally obtain the total acceleration

$$\langle \mathbf{a}(t) \rangle = \int \frac{p^2/2m - \epsilon_0}{\Delta_s} \frac{1}{m} \langle \chi_0(t) | (-\nabla V_C) | \mathbf{P} \rangle |\mathcal{F}_s [p'_z(t)]|^2 \times \left\{ -\frac{i}{\hbar} \int_{t_0}^t e^{i\frac{1}{\hbar} S(\mathbf{P}, t, t_1)} e^{-i\frac{\epsilon_0}{\hbar}(t_1 - t_0)} V_F(\mathbf{P}, t_1) \eta(\mathbf{P}, t_1) dt_1 \right\} d\mathbf{P}. \quad (\text{A21})$$

## Appendix B. Tunnel of a distant wavepacket

In this appendix, we compute the tunnel probability of a wavepacket located at a finite distance from the barrier's edge. In a description using localized atomic orbitals (LAO), the molecular orbital is decomposed into a basis of wavepackets, each centered at a different ion site. In molecules at equilibrium nuclear distances, these localized wavefunctions are generally not separated by internal potential barriers, therefore



dwelling in a multi-ion Coulomb potential well. The presence of a strong field modulates the molecular potential, forming an external potential barrier, that separates the bound orbitals from the continuum. The extended nature of the molecular well, therefore, results in some of the LAO positioned at finite, non zero, distances from the edge of the barrier (see figure 1 in the main text).

Let us consider a stationary wavepacket with energy  $\epsilon_0$  centered at the coordinate  $z_c$ , at the left of a potential barrier  $V(z)$ . The barrier's edges  $z_b$  and  $z_f$  are defined so that  $V(z) = 0$  if  $z < z_b$ , and  $V(z) > \epsilon_0$  if  $z_b < z < z_f$  (see figure 10).

We express the wavepacket as a planewave decomposition,  $\chi(z, t) = e^{-i\epsilon_0 t/\hbar} \chi_0(z)$ , with

$$\chi_0(z) = \int_{-\infty}^{\infty} \tilde{g}(p) e^{iS(p,z)/\hbar} e^{ip(z-z_c)/\hbar} dp \simeq \int_{-\infty}^{\infty} \tilde{g}(p) e^{iS_{\text{WKB}}(z)/\hbar} e^{-ipz_c/\hbar} dp, \quad (\text{B1})$$

where we have defined  $S_{\text{WKB}}(z) = \int_{-\infty}^z \sqrt{2m[\epsilon - V(z')]} dz'$  as the semiclassical approximation ( $\hbar \rightarrow 0$ ) to the planewave phase  $S(p, z) + pz$ , according to the 0th-order WKB approximation [62]. Note that it is crucial to preserve the localized nature of the wavepacket, therefore the 0th order WKB approximation is used at the level of the individual planewaves composing it, rather than to the total wavefunction. We can rewrite equation (B1) as

$$\chi(z, t) = e^{-i\epsilon_0 t/\hbar} e^{iS_{\text{WKB}}(z)/\hbar} \int_{-\infty}^{\infty} \tilde{g}(p) e^{-ipz_c/\hbar} dp. \quad (\text{B2})$$

Defining  $S_{\text{WKB}}(z_b)$  as the semiclassical approximation to  $S(p, z_b) + pz_b$ , equation (B2) can be approximated by

$$\chi_0(z) \simeq e^{i[S_{\text{WKB}}(z) - S_{\text{WKB}}(z_b)]/\hbar} \int_{-\infty}^{\infty} \tilde{g}(p) e^{iS(p, z_b)} e^{ip(z_b - z_c)/\hbar} dp \quad (\text{B3})$$

$$= e^{i[S_{\text{WKB}}(z) - S_{\text{WKB}}(z_b)]/\hbar} \chi_0(z_b). \quad (\text{B4})$$

The tunnel transmission probability of the particle located at the wavepacket's mean position  $z_c$ , is therefore given by

$$P(z_c, z_b) = \frac{|\chi_0(z_f)|^2}{|\chi_0(z_c)|^2} = W_{\text{WKB}} \frac{|\chi_0(z_b)|^2}{|\chi_0(z_c)|^2}, \quad (\text{B5})$$

where  $W_{\text{WKB}}$  corresponds to the WKB probability of tunneling of a particle located at the left edge of the barrier [62],

$$W_{\text{WKB}} = e^{-\frac{2}{\hbar} \int_{z_b}^{z_f} \sqrt{2m[V(z') - \epsilon_0]} dz'}. \quad (\text{B6})$$

Therefore, equation (B5) can be interpreted as the probability of tunneling times the probability of the particle being near the edge of the barrier. The quotient in equation (B5) shows that the tunnel probability is reduced when the wavepacket's center is at a distance  $z_b - z_c$  from the edge of the barrier. We define the ratio of the tunnel amplitude probability of a distant wavepacket to the amplitude probability of the particle being at the barrier's edge is given by

$$T(z_c, z_b) = \left[ \frac{P(z_c, z_b)}{P(z_b, z_b)} \right]^{1/2} = \frac{|\chi_0(z_b)|}{|\chi_0(z_c)|}. \quad (\text{B7})$$

In the case of the  $\text{H}_2^+$  molecule, the wavepackets  $\chi_0$  are atomic orbitals localized at the molecular ion sites,  $z_c = \pm R/2$ , where  $R$  is the internuclear distance. The coordinate of the barrier's edge,  $z_b$ , corresponds approximately to  $-R/2$  ( $R/2$ ) for a positive (negative) field amplitude. To introduce the site-dependent probabilities in the SFA formalism it is sufficient to redefine the molecular form factor,  $\mathcal{F}_s [p'_z(t)]$ , in the

s-SFA formula (A21) as

$$\mathcal{F}_{\text{SDT}}^{\pm} [p'_z(t)] = T \left( -\frac{R}{2}, \pm \frac{R}{2} \right) e^{i\frac{p'_z(t)R}{\hbar}} + T \left( \frac{R}{2}, \pm \frac{R}{2} \right) e^{-i\frac{p'_z(t)R}{\hbar}}, \quad (\text{B8})$$

where the positive (negative) index corresponds to ionization through a barrier at the right (left) side of the molecule. Note however that the modulus of the form factor is the same, regardless the right/left position of the barrier, and therefore we can drop the superindex,

$$|\mathcal{F}_{\text{SDT}} [p'_z(t)]|^2 = |\mathcal{F}_{\text{SDT}}^+ [p'_z(t)]|^2 = |\mathcal{F}_{\text{SDT}}^- [p'_z(t)]|^2. \quad (\text{B9})$$

Therefore equation (A21) finally becomes

$$\begin{aligned} \langle \mathbf{a}(t) \rangle &= \int \frac{p'/2m - \epsilon_0}{\Delta_s} \frac{1}{m} \langle \chi_0 | (-\nabla V_C) | \mathbf{P} \rangle |\mathcal{F}_{\text{SDT}} [p'_z(t)]|^2 e^{i\frac{\epsilon_0}{\hbar}(t-t_0)} \\ &\times \left\{ -\frac{i}{\hbar} \int_{t_0}^t e^{i\frac{1}{\hbar}S(\mathbf{P},t,t_1)} e^{-i\frac{\epsilon_0}{\hbar}(t_1-t_0)} V_F(\mathbf{P}, t_1) \langle \mathbf{P} | \frac{C_F}{r^n} | \chi_0 \rangle dt_1 \right\} d\mathbf{P}. \end{aligned} \quad (\text{B10})$$

The new form factor,  $\mathcal{F}_{\text{SDT}}$ , defines the difference between the standard SFA (s-SFA) and the site-dependent-tunneling SFA (SDT-SFA) proposed in this work.

## ORCID iDs

Laura Rego  <https://orcid.org/0000-0002-7214-7402>

Carlos Hernández-García  <https://orcid.org/0000-0002-6153-2647>

Antonio Picón  <https://orcid.org/0000-0002-6142-3440>

Luis Plaja  <https://orcid.org/0000-0001-8709-7295>

## References

- [1] Corkum P B 1993 Plasma perspective on strong field multiphoton ionization *Phys. Rev. Lett.* **71** 1994
- [2] Schafer K J, Yang B, DiMauro L F and Kulander K C 1993 Above threshold ionization beyond the high harmonic cutoff *Phys. Rev. Lett.* **70** 1599
- [3] Peng P, Marceau C and Villeneuve D 2019 Attosecond imaging of molecules using high harmonic spectroscopy *Nat. Rev. Phys.* **1** 144
- [4] Lein M, Hay N, Velotta R, Marangos J P and Knight P L 2002 Interference effects in high-order harmonic generation with molecules *Phys. Rev. A* **66** 023805
- [5] Lein M, Hay N, Velotta R, Marangos J P and Knight P L 2002 Role of the intramolecular phase in high-harmonic generation *Phys. Rev. Lett.* **88** 183903
- [6] Lein M, Corso P P, Marangos J P and Knight P L 2003 Orientation dependence of high-order harmonic generation in molecules *Phys. Rev. A* **67** 023819
- [7] Baker S et al 2008 Dynamic two-center interference in high-order harmonic generation from molecules with attosecond nuclear motion *Phys. Rev. Lett.* **101** 053901
- [8] Haessler S et al 2010 Attosecond imaging of molecular electronic wavepackets *Nat. Phys.* **6** 200206
- [9] Yun H, Lee K-M, Sung J H, Kim K T, Kim H T and Nam C H 2015 Resolving multiple molecular orbitals using two-dimensional high-harmonic spectroscopy *Phys. Rev. Lett.* **114** 153901
- [10] Luu T T, Garg M, Kruchinin S Y, Moulet A, Hassan M T and Goulielmakis E 2015 Extreme ultraviolet high-harmonic spectroscopy of solids *Nature* **521** 498
- [11] Silva R E F, Blinov I V, Rubtsov A N, Smirnova O and Ivanov M 2018 High-harmonic spectroscopy of ultrafast many-body dynamics in strongly correlated systems *Nat. Photon.* **12** 266
- [12] Landsman A S, Weger M, Maurer J, Boge R, Ludwig A, Heuser S, Cirelli C, Gallmann L and Keller U 2014 Ultrafast resolution of tunneling delay time *Optica* **1** 343–9
- [13] Itatani J, Levesque J, Zeidler D, Niikura H, Pépin H, Kieffer J C, Corkum P B and Villeneuve D M 2004 Tomographic imaging of molecular orbitals *Nature* **432** 867871
- [14] Patchkovskii S, Zhao Z, Brabec T and Villeneuve D M 2006 High harmonic generation and molecular orbital tomography in multielectron systems: Beyond the single-active-electron approximation *Phys. Rev. Lett.* **97** 123003
- [15] Ciriolo A G, Devetta M, Faccialà D, Geetha P P, Pusala A, Vozzi C and Stagira S 2018 Molecular orbital tomography based on high-order harmonic generation: principles and perspectives *Advances in the Application of Lasers in Materials Science* ed P M Ossi (Berlin: Springer) 274 pp 191–216
- [16] Keldysh L V 1965 Ionization in the field of a strong electromagnetic wave *Zh. Eksp. Teor. Fiz.* **20** 1307
- [17] Faisal F H M 1973 Multiple absorption of laser photons by atoms *J. Phys. B: At. Mol. Phys.* **6** L89
- [18] Reiss H R 1980 Effect of an intense electromagnetic field on a weakly bound system *Phys. Rev. A* **22** 1786
- [19] Lewenstein M, Balcou P, Ivanov M Y, L'Huillier A and Corkum P B 1994 Theory of high-harmonic generation by low-frequency laser fields *Phys. Rev. A* **49** 2117
- [20] Salières P et al 2001 Feynman's path-integral approach for intense-laser-atom interactions *Science* **292** 902
- [21] Pérez-Hernández J A, Roso L and Plaja L 2009 Harmonic generation beyond the strong-field approximation: the physics behind the short-wave-infrared scaling laws *Opt. Express* **17** 9891

- [22] Chirilă C C and Lein M 2006 Strong-field approximation for harmonic generation in diatomic molecules *Phys. Rev. A* **73** 023410
- [23] Zhou X X, Tong X M, Zhao Z X and Lin C D 2005 Alignment dependence of high-order harmonic generation from and molecules in intense laser fields *Phys. Rev. A* **72** 033412
- [24] Zimmermann B, Lein M and Rost J M 2005 Analysis of recombination in high-order harmonic generation in molecules *Phys. Rev. A* **71** 033401
- [25] Lagmago Kamta G and Bandrauk A D 2005 Three-dimensional time-profile analysis of high-order harmonic generation in molecules: Nuclear interferences in  $H_2^+$  *Phys. Rev. A* **71** 053407
- [26] Vozzi C, Negro M, Calegari F, Sansone G, Nisoli M, De Silvestri S and Stagira S 2011 Generalized molecular orbital tomography *Nat. Phys.* **7** 822
- [27] Miller M R, Jarón-Becker A and Becker A 2016 High-harmonic spectroscopy of laser-driven nonadiabatic electron dynamics in the hydrogen molecular ion *Phys. Rev. A* **93** 013406
- [28] Smirnova O, Mairesse Y, Patchkovskii S, Dudovich N, Villeneuve D, Corkum P and Yu M 2009 Ivanov, High harmonic interferometry of multi-electron dynamics in molecules *Nature* **460** 972–7
- [29] Baker S, Robinson J S, Haworth C A, Teng H, Smith R A, Chirilă C C, Lein M, Tisch J W G and Marangos J P 2006 Probing proton dynamics in molecules on an attosecond time scale *Science* **312** 424–7
- [30] Wagner N, Wuest A, Christov I, Popmintchev T, Zhou X, Murnane M and Kapteyn H 2006 Monitoring molecular dynamics using coherent electrons from high harmonic generation *Proc. Natl Acad. Sci. USA* **103** 13279
- [31] Blaga C I, Xu J, DiChiara A D, Sistrunk E, Zhang K, Agostini P, Miller T A, DiMauro L F and Lin C D 2012 Imaging ultrafast molecular dynamics with laser-induced electron diffraction *Nature* **483** 194197
- [32] Perelomov A M, Popov V S and Terentev M V 1966 Ionization of atoms in an alternating electric field *Zh. Eksp. Teor. Fiz.* **50** 1393
- [33] Krainov V P 1997 Ionization rates and energy and angular distributions at the barrier-suppression ionization of complex atoms and atomic ions *J. Opt. Soc. Am. B.* **14** 425–31
- [34] Ammosov M V, Delone N B and Krainov V P 1986 Tunnel ionization of complex atoms and of atomic ions in an alternating electromagnetic field *Zh. Eksp. Teor. Fiz.* **91**
- [35] Tong X M, Zhao Z X and Lin C D 2002 Theory of molecular tunneling ionization *Phys. Rev. A* **66** 033402
- [36] Muth-Böhm J, Becker A and Faisal F H M 2000 Suppressed molecular ionization for a class of diatomics in intense femtosecond laser fields *Phys. Rev. Lett.* **85** 2280
- [37] Tolstikhin O I, Morishita T and Madsen L B 2011 Theory of tunneling ionization of molecules: weak-field asymptotics including dipole effects *Phys. Rev. A* **84** 053423
- [38] Murray R, Spanner M, Patchkovskii S and Ivanov M Y 2011 Tunnel ionization of molecules and orbital imaging *Phys. Rev. Lett.* **106** 173001
- [39] Liu M and Liu Y 2016 Application of the partial-Fourier-transform approach for tunnel ionization of molecules *Phys. Rev. A* **93** 043426
- [40] Labeye M, Risoud F, Lévêque C, Caillaud J, Maquet A, Shaaran T, Salières P and Taieb R 2019 Dynamical distortions of structural signatures in molecular high-order harmonic spectroscopy *Phys. Rev. A* **99** 013412
- [41] Zuo T, Chelkowski S and Bandrauk A D 1993 Harmonic generation by the  $H_2^+$  molecular ion in intense laser fields *Phys. Rev. A* **48** 3837
- [42] Hay N, Velotta R, Lein M, de Nalda R, Heesel E, Castillejo M and Marangos J P 2002 High-order harmonic generation in laser-aligned molecules *Phys. Rev. A* **65** 053805
- [43] Lara-Astiaso M, Silva R E F, Gubaydullin A, Rivière P, Meier C and Martín F 2016 Enhancing high-order harmonic generation in light molecules by using chirped pulses *Phys. Rev. Lett.* **117** 093003
- [44] Hehre W J, Stewart R F and Pople J A 1969 self-consistent molecular-orbital methods. I. use of gaussian expansions of Slater-type atomic orbitals *J. Chem. Phys.* **51** 2657
- [45] Palke W E and Lipscomb W N 1966 Molecular SCF Calculations on  $CH_4$ ,  $C_2H_2$ ,  $C_2H_4$ ,  $C_2H_6$ ,  $BH_3$ ,  $B_2H_6$ ,  $NH_3$ , and  $HCN$  *J. Am. Chem. Soc.* **88** 2384–93
- [46] Smirnova O, Spanner M and Ivanov M 2007 Anatomy of strong field ionization II: to dress or not to dress? *J. Mod. Opt.* **54** 1019–38
- [47] Volkov D M 1935 Concerning a class of solution of the Dirac equation *Z. Phys.* **94** 250–60
- [48] Ciappina M F, Chirilă C C and Lein M 2007 Influence of Coulomb continuum wave functions in the description of high-order harmonic generation with  $H_2^+$  *Phys. Rev. A* **75** 043405
- [49] Suárez N, Chacón A, Pérez-Hernández J A, Biegert J, Lewenstein M and Ciappina M F 2017 High-order-harmonic generation in atomic and molecular systems *Phys. Rev. A* **95** 033415
- [50] Anh-Thu. L, Della Picca R, Fainstein P D, Telnov D A, Lein M and Lin C D 2008 Theory of high-order harmonic generation from molecules by intense laser pulses *J. Phys. B: At. Mol. Opt. Phys.* **41** 081002
- [51] Zair A et al 2008 Quantum path interferences in high-order harmonic generation *Phys. Rev. Lett.* **100** 143902
- [52] Lewenstein M, Salières P and L’Huillier A 1995 Phase of the atomic polarization in high-harmonic generation *Phys. Rev. A* **52** 4747
- [53] Kling M F et al 2006 Control of electron localization in molecular dissociation *Science* **312** 246
- [54] He F, Ruiz C and Becker A 2007 Control of electron excitation and localization in the dissociation of  $H_2^+$  and its isotopes using two sequential ultrashort laser pulses *Phys. Rev. Lett.* **99** 083002
- [55] Sansone G et al 2010 Electron localization following attosecond molecular photoionization *Nature* **465** 763
- [56] Gaarde M B, Tate J L and Schafer K J 2008 Macroscopic aspects of attosecond pulse generation *J. Phys. B: At. Mol. Opt. Phys.* **41** 13
- [57] Rundquist A, Durfee C, Chang Z, Herne C, Kapteyn H and Murnane M 1998 Phase-matched generation of coherent soft x-rays *Science* **280** 14121415
- [58] Hernández-García C, Pérez-Hernández J A, Ramos J, Jarque E C, Roso L and Plaja L 2010 High-order harmonic propagation in gases within the discrete dipole approximation *Phys. Rev. A* **82** 033432
- [59] Suárez N, Chacón A, Ciappina M F, Wolter B, Biegert J and Lewenstein M 2016 Above-threshold ionization and laser-induced electron diffraction in diatomic molecules *Phys. Rev. A* **94** 043423
- [60] Becker A, Plaja L, Moreno P, Nurhuda M and Faisal F H M 2001 Total ionization rates and ion yields of atoms at nonperturbative laser intensities *Phys. Rev. A* **64** 023408
- [61] Pérez-Hernández J A, Ramos J, Roso L and Plaja L 2010 Harmonic generation beyond the strong-field approximation: phase and temporal description *Laser Phys.* **20** 1044–50
- [62] Schiff L I 1949 *Quantum Mechanics* (New York: McGraw-Hill)