Ab-initio angle and energy resolved photoelectron spectroscopy with
time-dependent density-functional theory

U. De Giovannini, D. Varsano, M. A. L. Marques, H. Appel, E. K. U. Gross, and A. Rubio
1) University of the Basque Country UPV/EHU, Nano-Bio Spectroscopy Group, Avenida de Tolosa 72, 20018 San
Sebastian, Spain
2) Department of Physics, University of Rome "La Sapienza", Piazzale Aldo Moro 5, 00185 Roma, Italy
3) Université de Lyon, F-69000 Lyon, France and LPMCN, CNRS, UMR 5586, Université Lyon 1, F-69622
Villeurbanne, France
4) Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, D-14195 Berlin, Germany
5) Max-Planck Institute for Microstructure Physics, Weinberg 2, 06120 Halle, Germany

(Dated: 2 May 2014)

We present a time-dependent density-functional method able to describe the photoelectron spectrum of atoms and
molecules when excited by laser pulses. This computationally feasible scheme is based on a geometrical
partitioning that efficiently gives access to photoelectron spectroscopy in time-dependent density-functional
calculations. By using a geometrical approach, we provide a simple description of momentum-resolved photo-
mission including multi-photon effects. The approach is validated by comparison with results in the literature
and exact calculations. Furthermore, we present numerical photoelectron angular distributions for randomly
oriented nitrogen molecules in a short near infrared intense laser pulse and helium-(I) angular spectra for
aligned carbon monoxide and benzene.

PACS numbers: 31.15.E-, 33.60.+q, 33.80.Eh, 33.80.Rv

I. INTRODUCTION

Photoelectron spectroscopy is a widely used tech-
nique to analyze the electronic structure of complex sys-
tems. The advent of intense ultra-short laser sources
has extended the range of applicability of this technique
to a vast variety of non-linear phenomena like high-
harmonic generation, above-threshold ionization (ATI),
bond softening and vibrational population trapping. Further-
more, it turned attosecond time-resolved pump-probe photoelectron spectroscopy into a powerful tech-
nique for the characterization of excited-states dynamics in nano-structures and biological systems. Angular-
resolved ultraviolet photoelectron spectroscopy is by
now established as a powerful technique for studying
geometrical and electronic properties of organic
thin films. Time-resolved information from streaking
spectrograms, shearing interferograms, photoelectron
diffraction, photoelectron holography, etc. hold the
promise of wavefunction reconstruction together with the
ability to follow the ultrafast dynamics of electronic
wavepackets. Clearly, to complement all these experi-
mental advances, and to help to interpret and understand
the wealth of new data, there is the need for ab-initio the-
ories able to provide (time-resolved) photoelectron spec-
tra (PES) and photoelectron angular distributions (PAD)
for increasingly complex atomic and molecular systems
subject to arbitrary perturbations (laser intensity and
shape).

Photoelectron spectroscopy is a general term which
refers to all experimental techniques based on the pho-
etronic effect. In photoemission experiments a light
beam is focused on a sample, transferring energy to the
electrons. For low light intensities an electron can ab-
sorb a single photon and escape from the sample with a
maximum kinetic energy $\hbar \omega - I_P$ (where $\omega$ is the photon angular frequency and $I_P$ the first ionization potential
of the system) while for high intensities electron dynam-
icss can be interpreted considering a three-step model. This model provides a semiclassical picture in terms of
ionization followed by free electron propagation in the
laser field with return to the parent ion, and rescattering.
Such rescattering processes are the source of many interesting physical phenomena. In the case of long pulses,
for instance, multiple photons can be absorbed resulting in
emerging kinetic energies of $s \hbar \omega - I_P - U_P$ (where $s$ is
the number of photons absorbed, $U_P = e^2 / 4 \omega^2$ is the ponderomotive energy and $e$ the electric field amplitude)
forming the so called ATI peaks in the resulting pho-
etronic spectrum. In all cases the observable is the
escaping electron momentum measured at the detector.

In general, the interaction between electrons in an
atom or molecule and a laser field is difficult to treat
theoretically, and several approximations are usually per-
formed. Clearly, a full many-body description of PES is
prohibitive, except for the case of few (one or two) electron
systems. As a consequence, the direct solution of the
time dependent Schrödinger equation (TDSE) in
the so-called single-active electron (SAE) approximation is a standard investigation tool for many strong-field ef-

a) Electronic mail: umberto.degiovannini@ehu.es

arXiv:1206.6031v1 [physics.atom-ph] 25 Jun 2012
for analytic and semi-analytic models. Perturbative approaches based on the standard Fermi golden rule are usually employed. For weak lasers, plane wave methods and the independent atomic center approximation have been applied, while in the strong field regime, Floquet theory, the strong-field approximation and semiclassical methods are routinely used. From a numerical point of view, it would be highly desirable to have a PES theory based on time-dependent density functional theory (TDDFT) where the complex many-body problem is described in terms of a fictitious single-electron system. For a given initial many body state, TDDFT maps the whole many-body problem into the time dependence of the density from which all physical properties can be obtained. The method is in principle exact, but in practice approximations have to be made for the unknown exchange-correlation functional as well as for specific density-functional providing physical observables. This latter issue is much less studied than the former, and to the best of our knowledge a formal derivation of momentum-resolved PES from the time dependent density has not been performed up to now. In any case, several works were published addressing the problem of single and multiple ionization processes within TDDFT. For example, ionization rates were calculated for atoms and molecules and TDDFT with the sampling point method (SPM) has been employed in the study of PES and PAD for sodium clusters. In this work, besides presenting a formal derivation of a photoelectron orbital functional, we report on a new and simple one-dimensional model also in comparison with the sampling point method. The above threshold ionization of three-dimensional hydrogen is examined and compared with values from the literature. In the second application we illustrate PADs from randomly oriented nitrogen molecules in a strong near-infrared ultra-short laser pulse. Comparison with the experimental and molecular strong-field approximation is discussed. The third application of the method regards helium-(I) (wavelength 58 nm) PADs for oriented carbon monoxide and benzene. Results are discussed in comparison with the plane wave approximation. Finally, in Sect. V, we discuss the results and present the conclusions.

All our numerical calculations were performed with the real-time, real-space TDDFT code Octopus freely available under the GNU public license. Atomic units are used throughout unless otherwise indicated.

II. MODELING PHOTOELECTRON SPECTRA

In order to put in perspective the results of the present Article, we will give a brief introduction on the status of the principal techniques available for ab-initio PES calculations. We start our description with the methods employed to study one-electron systems. For one-electron systems PES can be calculated exactly from the direct solution of the TDSE. Several methods have been employed to extract PES information from the solution of the TDSE. The most direct and intuitive way is via direct projection methods where the PES is obtained by projecting the wave function at the end of the pulse onto the eigenstates describing the continuum. These eigenstates are extracted through the direct diagonalization of the Hamiltonian without including the interaction with the field. The momentum probability distribution can then be easily obtained from the Fourier transform of the continuum part of the time-dependent wavefunction.

Another approach, that avoids the calculation of the full continuum spectrum, involves the analysis of the exact wavefunction |Ψ⟩ after the laser pulse via a resolvent technique. In this case, the energy resolved PES is given by the direct projection on out-going wavefunctions with \( P(E) = \langle \Phi(E) | |\Psi⟩|^2 = \langle \Phi(E) | |\hat{D}(E) | |\Psi⟩⟩\), where \( \Phi(E) \) denotes an out-going (unbound) electron of energy \( E \) of the laser-free Hamiltonian, and \( |\hat{D}(E) | |\Psi⟩⟩ \) is the corresponding projection operator that can be conveniently approximated.

Normally, one needs accurate wave functions in a large space domain to obtain the correct distribution of the
ejected electrons. This is because the unbound parts of the wave packet spread out of the core region, and conventional expressions for the transition amplitude need these parts of the wave function. Solving the TDSE within all the required volume in space can easily become a very difficult computational problem. Several techniques were developed during the years to solve the problem. For simple cases these difficulties can be overcome by the use of spherical coordinates. Geometrical splitting techniques have also been employed. Furthermore, formulations in the Kramers-Henneberger frame of reference, and in momentum-space led to calculations with remarkable high precision. Recently a promising surface flux method has also been proposed.

The exact solution of the TDSE in three dimensions for more than two electrons is unfeasible and the limit rises to four electrons for one-dimensional models. Due to this limitation basically all ab-initio calculations for multi-electron systems are preformed under the SAE approximation. In the SAE only one electron interacts with the external field while the other electrons are frozen, and the TDSE is thus solved only for the active electron. This approximation has been successfully employed in several photoemission studies for atoms and molecules in strong laser fields. However, the failure of this simple model to describe multi-electron (correlation) effects calls for better schemes.

The inclusion of exchange-correlation effects for a system of many interacting electrons can be achieved within TDDFT while keeping the simplicity of working with a set of time-dependent (fictitious) single-particle orbitals. In spite of transferring all the many-body problem into an unknown exchange-correlation functional, the lack of a density functional providing the electron emission probability is a major limitation for a direct access to photoelectron observables from the time evolution of the density (note that, in spite of the Runge-Gross theorem, stating that all observables are functionals of the time dependent density, in practice we know few observables that can be written in terms of the time-dependent density, one example being the absorption spectra).

There have been some attempts to describe PES and multiple ionization processes with TDDFT in the standard adiabatic approximation. All these works use boundary absorbers to separate the bound and continuum part of the many-body wavefunction. The emission probability is then correlated with the time dependence of the number of bound electrons.

An alternative and simple scheme is provided by the SPM. Here the idea is to record single-particle wavefunctions in time at a fixed sampling point \( r_S \) away from the core. The time Fourier transform of the wavefunction recorded at \( r_S \) represents the probability of having an electron in \( r_S \) with energy \( E \). The probability to detect one electron with energy \( E \) in \( r_S \) is then given by the sum over all occupied orbitals:

\[
P_{r_S}(E) = \sum_{i=1}^{\text{occ.}} |\psi_i(r_S,E)|^2.
\]  

This method is easy to implement, can be extended to give also angular information, and is also clearly applicable to the TDSE in the SAE. However, it lacks formal derivation as it is directly based on Kohn-Sham wavefunctions without a direct connection to the many-body state. Furthermore, it is strongly dependent on the position of the sampling point and the minimum distance. This distance sometimes turns out to be quite large in order to avoid artifacts, and is strongly dependent on the laser pulse properties. We discuss further details concerning this method in Sect. [III A].

In the following we present an alternative method inspired by geometrical splitting and derive it from a phase-space point of view. The method can be naturally converged by increasing the size of the different simulation boxes.

### A. Phase-space geometrical interpretation

An intuitive description of photoelectron experiments can be obtained resorting to a phase-space picture. Experimental detectors are able to measure photoelectron velocity with a certain angular distribution for a sequence of ionization processes with similar initial conditions. The quantity available at the detector is therefore connected to the probability to register an electron with a given momentum \( \mathbf{p} \) at a certain position \( \mathbf{r} \). From this consideration it would be tempting to interpret photoemission experiments with a joint probability distribution in the phase-space \((\mathbf{r}, \mathbf{p})\). Such a classical picture however conflicts with the fundamental quantum mechanics notion of the impossibility to simultaneously measure momentum and position, and prevents us from proceeding in this direction. A link between the classical and quantum picture is needed beforehand.

In order to make a connection to a microscopic description it turns out to be convenient to extend the classical concept of phase-space distributions to the quantum realm. A common prescription comes from the Wigner transform of the one-body density matrix with respect to the center of mass \( \mathbf{R} = (\mathbf{r} + \mathbf{r}')/2 \) and relative \( \mathbf{s} = \mathbf{r} - \mathbf{r}' \) coordinates. The \( d \)-dimensional (here and after \( d \leq 3 \)) transform is defined as

\[
w(\mathbf{R}, \mathbf{p}, t) = \int \frac{d\mathbf{s}}{(2\pi)^d} e^{i\mathbf{p}\cdot\mathbf{s}} \rho(\mathbf{R} + \mathbf{s}/2, \mathbf{R} - \mathbf{s}/2, t),
\]  

with

\[
\rho(\mathbf{r}, \mathbf{r}', t) = \int d\mathbf{r}_2 \ldots d\mathbf{r}_N \Psi(\mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_N, t)
\]

\[
\times \Psi^*(\mathbf{r}', \mathbf{r}_2, \ldots, \mathbf{r}_N, t),
\]  

where

\[
\Psi(\mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_N, t) = \sum_{\text{occ.}} \psi_i(\mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_N, t).
\]
be the one-body density matrix, and
\( \Psi(r_1, r_2, \ldots, r_N, t) \) the N-body wavefunction of the
system at time \( t \). The Wigner function defined above is
normalized and its integral over the whole space (momentum) gives the probability to find an electron
with momentum \( p \) (position \( \mathbf{R} \)). As the uncertainty
principle prevents the simultaneous knowledge of po-
sition and momentum, \( w(\mathbf{R}, p) \) cannot be a proper
joint distribution. Moreover it can assume negative
values due to nonclassical dynamics. Nevertheless the
Wigner function \( w(\mathbf{R}, p) \) constitutes a concept close to a
probability distribution in phase space (\( \mathbf{R}, p \)) compatible
with quantum mechanics.

The quantum phase-space naturally leads to a geomet-
rical interpretation of photoemission. One could think to
divide the space in two regions \( A \) and \( B \) as in Fig. 1(a),
where region \( B \) represents the region where detectors are
positioned and \( A \) is defined as the complement of \( B \). In
this picture, PES can be seen as the probability to have
an electron with given momentum in \( B \). It is then natu-
ral to define the momentum-resolved photoelectron spectrum as

\[
P(p) = \lim_{t \to \infty} \int_B d\mathbf{R} w(\mathbf{R}, p, t) \tag{4}
\]

where the spatial integration is carried out in region \( B \),
and the limit \( t \to \infty \) assures that region \( B \) contains all
photoelectron contributions. From the knowledge of the
momentum-resolved PES [cf. Eq. (3)] one can access sev-
eral different quantities by simple integration. For in-
stance, in three dimensions \( (d = 3) \) the energy-resolved
PES is obtained integrating over the solid angle \( \Omega \):

\[
P(E = p^2/2) = \int_0^{4\pi} d\Omega P(p) \,. \tag{5}
\]

and the photoelectron angular distribution in the system
reference frame is given by,

\[
P(E = p^2/2, \theta) = \int_0^{2\pi} d\phi P(p) \,. \tag{6}
\]

In spite of giving an intuitive picture of PES, Eq. (4)
is not suited for direct numerical evaluation since it re-
quires the knowledge of the full one-body density matrix
in the whole space. In the next section we will make a
contact with effective single particle theories like TDDFT
to overcome the limitations due to the knowledge of the
many-body wavefunction. In order to avoid integration
over the whole space an efficient evolution scheme is pre-
sented in Sect. III C.

B. Phase space interpretation within TDDFT

TDDFT is an effective single particle theory where
the many-body wavefunction is described by an auxil-
iary single Slater determinant \( \Psi_{KS}(r_1, \ldots, r_N) \) built out
of Kohn-Sham orbitals \( \psi_i(r) \). In order to simplify the
notation, we drop the explicit time dependence from the
wavefunctions and assume that the following equations
are written in the limit \( t \to \infty \) as prescribed by Eq. (4).

Being represented by a single determinant, the one-
body Kohn-Sham density matrix is given by

\[
\rho_{KS}(\mathbf{r}, \mathbf{r}') = \sum_{i=1}^{\text{occ.}} \psi_i(\mathbf{r}) \psi_i(\mathbf{r}') \tag{7}
\]

where the sum in carried out over all occupied orbitals.
Performing a decomposition of each orbital according to
the partition of Fig. 1(a) we obtain

\[
\psi_i(\mathbf{r}) = \psi_{A,i}(\mathbf{r}) + \psi_{B,i}(\mathbf{r}) \,, \tag{8}
\]

where \( \psi_{A,i}(\mathbf{r}) \) is the part of the wavefunction describ-
ing states localized in \( A \) and \( \psi_{B,i}(\mathbf{r}) \) is the ionized con-
tribution measured at the detector in \( B \). The one-body
density matrix can now be accordingly decomposed as a
sum of four terms

\[
\rho_{KS}(\mathbf{r}, \mathbf{r}') = \sum_{i=1}^{\text{occ.}} \left[ \psi_{A,i}(\mathbf{r}) \psi_{A,i}^*(\mathbf{r}') + \psi_{A,i}(\mathbf{r}) \psi_{B,i}^*(\mathbf{r}') \right.
\]

\[
+ \left. \psi_{B,i}(\mathbf{r}) \psi_{A,i}^*(\mathbf{r}') + \psi_{B,i}(\mathbf{r}) \psi_{B,i}^*(\mathbf{r}') \right] \,. \tag{9}
\]

From Eq. (9) we can build the KS Wigner function de-
defined in Eq. (4) and obtain the momentum-resolved prob-
ability distribution by inserting it into Eq. (4). We
note that this step involves a non-trivial approximation,
that is, the KS one-body density matrix is a good
approximation to the fully interacting one in region \( B \).
This is, however, much milder than the assumption that
the Kohn-Sham determinant is a good approximation to
the many-body wavefunction in region \( B \), as it is done,
e.g., in the SPM.

The final result is a sum of four overlap double inte-
grals that can be simplified further. For a detailed calcu-
lation we refer to Appendix A. The first overlap integral,
containing a product of two functions localized in \( A \) [cf.
Eq. (9)], is zero due to the spatial integration in \( B \). The
two next overlap integrals, containing mixed products of
wavefunctions localized in $A$ and $B$, can be reduced by increasing the size of region $A$. Assuming $A$ to be large enough to render these terms negligible the only integral we are left with is the one containing functions in $B$, leading to

$$P(p) \approx \int_B dR \int \frac{ds}{(2\pi)^2} e^{ipt} \sum_{i=1}^{\text{occ.}} \psi_{B,i}(R + \frac{s}{2}) \psi_{B,i}^*(R - \frac{s}{2}).$$

(10)

The approximation sign $\approx$ is a reminder for the error committed in discarding the mixed overlap integrals. Since the probability of finding an ionized electron in region $A$ is zero for $t \to \infty$, we can extend the integration over $B$ in Eq. (10) to the whole space. Using the integral properties of the Wigner transform we finally obtain

$$P(p) \approx \sum_{i=1}^{\text{occ.}} |\tilde{\psi}_{B,i}(p)|^2,$$

(11)

where $\tilde{\psi}_{B,i}(p)$ is the Fourier transform of $\psi_{B,i}(r)$ and the expression is written in the limit for $t \to \infty$. Equation (11) gives an intuitive formulation of momentum-resolved PES as a sum of the Fourier component of each orbital in the detector region. It is worth to note that Eq. (11) is not restricted to TDDFT and can be applied to other effective single-particle formulations such as time-dependent Hartree-Fock and the TDSE in the SAE approximation.

The numerical evaluation of the ionization probability from Eq. (11) requires the knowledge of the wavefunction after the external field has been switched off. For ionization processes this means that one has to deal with simulation boxes that extend over several hundred atomic units and this practically constrains the method only to one-dimensional calculations. In the next section we will derive a simple scheme to overcome this limitation making the present scheme applicable for realistic simulations of molecules and nanostructures.

C. The mask method

In the previous sections we described a practical way to evaluate the momentum-resolved PES following the spatial partitioning of Fig. 1 (a) and how this can be conveniently cast in the language of TDDFT. In this section we take a step further in developing an efficient time evolution scheme by exploiting the geometry of the problem together with some physical assumptions.

We start by introducing a split-evolution scheme: At each time $t$ we implement a spatial partitioning of Eq. (8) as following

$$\begin{cases} 
\psi_{A,i}(r,t) = M(r)\psi_i(r,t) \\
\psi_{B,i}(r,t) = [1 - M(r)]\psi_i(r,t)
\end{cases},$$

(12)

where $M(r)$ is a smooth mask function defined to be 1 deep in the interior of region $A$ and 0 outside, as shown in Fig. 2. Such a mask function, along with the partitions $A$ and $B$, introduces a buffer region $C$ (technically handled as the outermost shell of $A$), where $\psi_{A,i}(r,t)$ and $\psi_{B,i}(r,t)$ overlap [see Fig. 1 (b)].

We can set up a propagation scheme from time $t$ to $t'$ as following

$$\begin{cases} 
\psi_{A,i}(r,t') = M(r)U(t',t) [\psi_{A,i}(r,t) + \psi_{B,i}(r,t)] \\
\psi_{B,i}(r,t') = [1 - M(r)]U(t',t) [\psi_{A,i}(r,t) + \psi_{B,i}(r,t)]
\end{cases},$$

(13)

where $U(t',t)$ is the time propagator associated with the full Hamiltonian including the external fields. Equation (13) defines a recursive propagation scheme completely equivalent to a time propagation in the whole space $A \cup B$.

In typical experimental setups, detectors are situated far away from the sample and electrons overcoming the ionization barrier travel a long way before being detected. During their journey toward the detector, and far away from the molecular system, they practically evolve as free particles driven by an external field. It seems therefore a waste of resources to solve the full Schrödinger equation for the traveling electrons while their behavior can be described analytically. In addition, an ideal detector placed relatively close to the molecular region would measure the same PES.

From these observations we conclude that we can reduce region $A$ to the size of the interaction region and assume electrons in $B$ to be well described by non-interacting Volkov states. Volkov states are the exact solution of the Schrödinger equation for free electrons in an oscillating field. They are plane-waves and are therefore naturally described in momentum space. In the velocity gauge the Volkov time propagator is formally expressed
by
\[ U_V(t', t) = \exp \left\{ -i \int_t^{t'} d\tau \frac{1}{2} \left[ \mathbf{p} - \frac{\mathbf{A}(\tau)}{c} \right]^2 \right\}, \] (14)

where the time-ordering operator is omitted for brevity and \( \mathbf{A}(\tau) \) is the vector potential. This is equivalent to the use of a strong-field approximation in the outer region in the same spirit of the Lewenstein model.\(^{20}\)

In summary, the method we propose consists in solving numerically the real-space TDDFT equations in \( A \) and analytically propagating the wavefunctions residing in \( B \) in momentum space. In this setup region \( C \) acts as a communication layer between functions in \( A \) and \( B \). Under this prescription, and by handling \( B \)-functions in momentum space, Eq. \((13)\) becomes

\[
\begin{align}
\psi_{A,i}(\mathbf{r}, t') &= \eta_{A,i}(\mathbf{r}, t') + \eta_{B,i}(\mathbf{r}, t') \\
\psi_{B,i}(\mathbf{p}, t') &= \xi_{A,i}(\mathbf{p}, t') + \xi_{B,i}(\mathbf{p}, t')
\end{align}
\] (15)

with

\[
\begin{align}
\eta_{A,i}(\mathbf{r}, t') &= M(\mathbf{r}) U(t', t) \psi_{A,i}(\mathbf{r}, t) \\
\eta_{B,i}(\mathbf{r}, t') &= M(\mathbf{r}) \int d\mathbf{p} e^{i \mathbf{p} \cdot \mathbf{r}} U_V(t', t) \tilde{\psi}_{B,i}(\mathbf{p}, t) \\
\xi_{A,i}(\mathbf{p}, t') &= \int d\mathbf{r} e^{-i \mathbf{p} \cdot \mathbf{r}} \left\{ 1 - M(\mathbf{r}) \right\} U(t', t) \psi_{A,i}(\mathbf{r}, t) \\
\xi_{B,i}(\mathbf{p}, t') &= U_V(t', t) \tilde{\psi}_{B,i}(\mathbf{p}, t) - \int d\mathbf{r} e^{-i \mathbf{p} \cdot \mathbf{r}} \eta_{B,i}(\mathbf{r}, t').
\end{align}
\] (16)

At each time step the orbital \( \psi_{A,i} \) is evolved under the mask function and stored in \( \eta_{A,i} \), forcing \( \eta_{A,i} \) to be localized in \( A \). At the same time, the components of \( \psi_{A,i} \) escaping from \( A \) are collected in momentum space by \( \xi_{A,i} \). We then add to \( \xi_{A,i} \) the contribution of the wavefunctions already present in \( B \) at time \( t \) by summing up \( U_V \tilde{\psi}_{B,i} \). In order to allow electrons to come back from \( B \) to \( A \) we include \( \eta_{B,i} \) in \( A \) and correct the function in \( B \) by removing its Fourier components [second term in Eq. \((16d)\)].

One of the advantages of Eq. \((15)\) is that all the spatial integrals present in \( \eta_{B,i}(\mathbf{r}, t') \) and \( \xi_{B,i}(\mathbf{p}, t') \) are performed on functions localized in \( C \). Therefore, integrals over the whole space are evaluated at the cost of an integration on the much smaller buffer region \( C \) that can be easily evaluated by fast Fourier transform algorithms. Similar considerations hold for integrals in momentum space under the assumption that \( B \)-functions \( \psi_{B,i}(\mathbf{p}, t) \) are localized in momentum. When region \( A \) is discretized on a grid, in order to avoid wavefunction wrapping at the boundaries and preserve numerical stability, additional care must be taken. In our implementation, numerical stability is addressed by the use of non-uniform Fourier transforms (see details in Appendix \[B\]).

There are situations were the electron flow from \( B \) to \( A \) is negligible. This is the case, for instance, when \( A \) is large enough to contain the whole wavefunctions at the time when the external field has been switched off. A propagation at later times will see photoelectrons flowing mainly from \( A \) to \( B \). In this situation, \( \eta_{B,i} \) and the corresponding correction term in \( \xi_{B,i} \) can be discarded. The evolution scheme of Eq. \((15)\) is thus simplified and becomes

\[
\begin{align}
\psi_{A,i}(\mathbf{r}, t') &= \eta_{A,i}(\mathbf{r}, t') \\
\psi_{B,i}(\mathbf{p}, t') &= \xi_{A,i}(\mathbf{p}, t') + \eta_{B,i}(\mathbf{r}, t') + U_V(t', t) \tilde{\psi}_{B,i}(\mathbf{p}, t)
\end{align}
\] (17)

In the following we will refer to Eq. \((17)\) as the “mask method” (MM), and to Eq. \((15)\) as the “full mask method” (FMM). We note here again that, being single-particle propagations schemes, both MM and FMM are not restricted to TDDFT and can be applied to other effective single-particle theories. As a matter of fact an approach similar to Eq. \((17)\) has already been employed in the propagation of the TDSE equations for atomic systems and in TDDFT for one-dimensional models of metal surfaces.\(^{27}\) We also note that the implementation of absorbing boundaries trough a mask function as done in Eq. \((16a)\) can be cast in terms of an additional imaginary potential (exterior complex scaling) in the Schrödinger equation. Such approach is commonly used in quantum optics.

Within the MM, the evolution in \( A \) is completely unaffected by the wavefunctions in \( B \) and ionized electrons are treated uniquely in momentum space. Compared with the FMM, the MM is numerically more stable as it is not affected by boundary wrapping. In order to achieve the conditions where Eq. \((17)\) is valid may require, however, large simulation boxes. Moreover as the mask function never absorbs perfectly the electrons, spurious reflections may appear. Suppression of such artifacts requires a further enlargement of the buffer region. With FMM spurious reflections are almost negligible.

Choosing between MM and FMM implies a tradeoff between computational complexity and numerical stability that strongly depends on the ionization dynamics of the process under study. In what follows, we will illustrate the differences and devise a prescription to help the choice of the most suitable method in each specific case.

### III. APPLICATIONS

In this section we present a few numerical applications of the schemes previously derived. In all calculations the boundary between \( A \) and \( B \) regions is chosen as a \( d \)-dimensional sphere implemented by the mask function:

\[
M(\mathbf{r}) = \begin{cases} 
1 & \text{if } r < R_C \\
1 - \sin^2 \left( \frac{(r - R_C)\pi}{2(R_C - R_A)} \right) & \text{if } R_C \leq r \leq R_A \\
0 & \text{if } r > R_A
\end{cases}
\] (18)
as shown in Fig. 2. Note that numerical studies (not presented here) revealed a weak dependence of the final results on the functional shape of the mask.

The time propagation of the orbitals in A is performed with the enforced time-reversal symmetry evolution operator\(^{39}\)

\[
U(t+\Delta t, t) = \exp \left( -i \frac{\Delta t}{2} H(t + \Delta t) \right) \exp \left( -i \frac{\Delta t}{2} H(t) \right),
\]

where \(H\) is the full KS Hamiltonian, and the coupling with the external field is expressed in the velocity gauge.

The first system we will study is hydrogen. In spite of being a one-electron system, it is a seemingly trivial case that has been and still is under thorough theoretical investigation\(^{23–25,49–52}\). Clearly we do not need TDDFT to study hydrogen and our numerical results are obtained by propagating the wavefunction with a non-interacting Hamiltonian. The interest in this case is focused on the numerical performance of different mask methods as hydrogen provides a useful benchmark.

The full TDDFT calculations performed for molecular nitrogen, carbon monoxide, and benzene are later presented in Sect. III B and Sect. III C respectively. In these cases, norm-conserving Troullier-Martins pseudopotentials and the exchange-correlation LB94 potential\(^{53}\) (that has the correct asymptotic limit for molecular systems) are employed. Finally, in all calculations the starting electronic structure of the molecules is calculated in the Born-Oppenheimer approximation at the experimental equilibrium geometry and the time evolution is performed with fixed ions.

### A. Photoelectron spectrum of hydrogen

As first example we study multi-photon ionization of a one-dimensional soft-core hydrogen atom, initially in the ground state, and exposed to a \(\lambda = 532\) nm (\(\omega = 0.0856\) a.u.) linearly polarized laser pulse with peak intensity \(I = 1.38 \times 10^{13}\) W/cm\(^2\), of the form

\[
A(t) = A_0 f(t) \cos(\omega t)
\]

where \(f(t)\) is a trapezoidal envelope function of 14 optical cycles with two-cycle linear ramps, constant for 10 cycles, and with \(A_0 = 31.7\) a.u. Here \(A(t)\) is the vector potential in units of the speed of light \(c\). A soft-Coulomb potential \(V(x) = -1/\sqrt{2 + x^2}\) is employed to model the electron-ion interaction. We propagate the electronic wavefunction in time and then compare the energy-resolved ionization probability obtained from different schemes. Along with MM, FMM, and SPM we present results for direct evaluation of PES from Eq. \(\text{[5]}\). In this method the spectrum is obtained by directly Fourier transforming the wavefunction in region B. Since the analysis is conducted without perturbing the evolution of the wavefunction we will refer to it as the “passive method” (PM).

This method requires the knowledge of the whole wavefunction after the pulse has been switched off, and since a considerable part of the wavepacket is far away from the core (for the present case a box of 500 a.u. radius is needed for 18 optical cycles), it is viable only for one-dimensional calculations. Nevertheless it is important as it constitutes the limiting case for both MM and FMM.

In Fig. 3 a color plot of the evolution of the electronic density as a function of time is shown. The electronic wavefunction splits into subpackets generated at each laser cycle (one optical cycle = \(2\pi/\omega = 1.774\) fs). These wavepackets evolve in bundles and their slope corresponds to a certain average momentum. ATI peaks are then formed by the build up of interfering wavepackets periodically emitted in the laser field and leading to a given final momentum\(^{12}\).

From Fig. 3 is it possible to see that electrons may be considered as escaped “already” at 30 a.u. away from the center. We set therefore \(R_A = 30\) a.u. and calculate energy-resolved PES with the PM. As we can see from Fig. 4 the spectrum presents several peaks at integer multiples of \(\omega\) following \(E = s\omega - I_P - U_P\) with \(U_P = A_0^2/4c^2 = 0.0133\) a.u. being the ponderomotive energy, \(I_P = 0.5\) a.u. the ionization potential, and \(s\) the number of absorbed photons. In this case the minimum number of photons needed to exceed the ionization threshold is \(s = 6\). Of course, the spectrum is only in qualitative agreement with three-dimensional calculation\(^{54}\) as expected from a one-dimensional soft-core model\(^{55}\).

PES calculated from MM, FMM, and SPM all agree as reported in Fig. 4. Numerical calculations were performed until convergence was achieved, leading to a grid with spacing \(\Delta R = 0.4\) a.u. and box sizes depending on the method. For MM we employed a simulation box of \(R_A = 70\) a.u. and set the buffer region at \(R_C = 30\) a.u.
In order to have energy resolution comparable with PM we used padding factors (see Appendix B) $P = P_N = 4$ and the total simulation time was $T = 18$ optical cycles. For FMM a smaller box of $R_A = 40$ a.u. with $R_C = 30$ a.u. is needed to converge results, and $P = 8$, $P_N = 2$ were needed to preserve numerical stability for $T = 18$ optical cycles. For SPM two sampling points at $r_S = -500, 500$ a.u. were needed to get converged results with a box of 550 a.u., and a complex absorber at 49 a.u. from the boundaries of the box. In addition, a total time of $T = 74$ optical cycles was required to collect all the wave packets. The need for such a huge box resides on the working conditions of SPM. In order to avoid spurious effects, the sampling points must be set at a distance such that the density front arrives after the external field has been switched off. Therefore the longer the pulse the further away the sampling points must be set. For these laser parameters one could rank each method according to increasing numerical cost starting from MM, followed by FMM, PM, and SPM.

As a second example we study the ionization of this one dimensional hydrogen atom by an ultra-short intense infrared laser. We employ a single two-cycle pulse of wavelength $\lambda = 800$ nm ($\omega = 0.057$ a.u.), intensity $I = 2.5 \times 10^{14}$ W/cm$^2$, and envelope

$$f(t) = \begin{cases} -\sin(\omega t/2N_c)^2 & \text{if } 0 \leq t \leq 2\pi N_c/\omega \\ 0 & \text{if } t > 2\pi N_c/\omega \end{cases}$$ (21)

with $N_c = 2$ and $A_0 = 225.8$ a.u.

Due to the laser strength and long wavelength, the electron evolution shown Fig. 5 is quite different from the one presented before. Electrons ejected from the core are driven by the laser and follow wide trajectories before returning to the parent ion. Such trajectories can be understood in the context of the semiclassical model where released electrons move as a free particle in a time-dependent field with a maximum oscillation amplitude of $x_0 = 2A_0/\omega c = 57.8$ a.u. Electrons ejected near a maximum of the electric field $\epsilon(t) = -\partial A(t)/\partial t$ are the ones gaining the most kinetic energy and are therefore responsible for the fast emerging electrons after rescattering with the core.

In Fig. 6 we show the energy-resolved PES for different methods. Here the spectra appear to be very far from any ATI structure due to short duration of the laser pulse and is characterized by some irregular maxima and minima.
The characteristic features of the ionization dynamics is strongly dependent on the detailed shape of the pulse as one can easily imagine by inspecting the asymmetry in the electron ejection from Fig. 5. Due to these dynamics, a dramatic carrier envelope phase dependence for such short pulses is expected.

All the different methods result in similar spectra but with different parameters. In PM we set $R_A = 50$ a.u. and a box of radius $R = 700$ a.u. is needed to contain the wave function after $T = 4$ optical cycles (one optical cycle $= 2.66$ fs). For MM $R_A = 200$ a.u., $R_C = 40$ a.u., and the padding factors are $P = 2$, $P_N = 4$. Here the value for $R_A$ is dictated by the width of the buffer region which needs to be wide enough to prevent spurious reflections. A considerably smaller box is needed for FMM, where $R_A = 60$ a.u., $R_C = 40$ a.u., $P = 4$, and $P_N = 2$. In this case one can reconstruct the total density in $A$ by evaluating $|\psi_A(x, t)|^2$ via Eq. (15) and compare it to the exact evolution. As one can see in Fig. 7, the reconstructed density displays a behavior remarkably similar to the exact one of Fig. 5, but with a considerably reduced computational cost. SPM requires sampling points at $r_S = -130, 130$ a.u. in a box of radius $R = 200$ a.u. with 49 a.u. wide complex adsorbers, and for a total time of $T = 7$ optical cycles.

The possibility to use relatively small simulation boxes is especially important for three-dimensional calculations where the computational cost scales with the third power of the box size. Both mask methods are practicable options for 3D simulations and the advantage of using FMM with respect to MM is driven by the electron dynamics. For long laser pulses MM appears to be more stable and it is a better choice than FMM, while for short pulses with large electron oscillations FMM can be more performant. SPM is a viable option for short pulses and small values for the oscillations.

As a last example, we present ATI of a real three-dimensional hydrogen atom subject to a long infrared pulse. We employ a laser linearly polarized along the $x$-axis with wavelength $\lambda = 800$ nm, intensity $I = 5 \times 10^{13}$ W/cm$^2$, pulse shape of the form (21) with $N_c = 20$ and $A_0 = 91.3$ a.u. Due to to the pulse length, MM appears to be the most appropriate choice in this case. In the calculation $R_A = 60$ a.u., $R_C = 50$ a.u., $P = 1$ and $P_N = 8$.

In Fig. 8, we show a high-resolution density plot of the PAD $P(E, \theta)$ defined in Eq. (16). The radial distance denotes the photoelectron energy while the angle indicates the direction of emission with respect to the laser polarization. The color density is plotted in logarithmic scale and represents the values of $P(E, \theta)$.

The photoelectron energy-angular distribution displays complex interference patterns. The pattern shape compares favorably with similar calculations in the literature.\textsuperscript{24,24,44,44,59} It consists of a series of rings with fine structures. Each ring represents the angular distribution of the photoelectron ATI peaks. The spacing of adjacent rings equals the photon energy $\omega = 0.057$ a.u. Photoelectrons are emitted mainly along the laser polarization, and the left-right symmetry of the rings indicates that the photoelectrons do not present any preferential ejection side with respect to the polarization axis. The first ring corresponds to the angular distribution of the first ATI peak. It presents a peculiar nodal pattern that is induced by the long-range Coulomb potential and is related to the fact that the ATI peak is determined by one dominant partial wave in the final state.\textsuperscript{24} The number of the stripes equals the angular momentum quantum number of the dominant partial wave in the final state plus one.\textsuperscript{24} In Fig. 8, the first ring contains six stripes and the dominant final state has angular momentum quantum number of 5. The pattern of the energy-angular distribution and the stripe number of the first ring are in good agreement with those in the literature.\textsuperscript{24,44} As for the fine structures, we observe that while the main ring pattern is already formed in the first half of the pulse, the fine structure builds up until the end of the pulse.
This supports the hypothesis that such structures are induced by the coherence of the two contributions from the leading and trailing edges of the pulse envelope\textsuperscript{13}

B. \textit{N}_2 under a few-cycle infrared laser pulse

In this section we compare theoretical and experimental angular resolved photoelectron probabilities for randomly oriented \textit{N}_2 molecules. We choose the laser parameters according to experiment\textsuperscript{28} i.e., we employ a \(N_e = 6\) cycle pulse of wavelength \(\lambda = 750\) nm (\(\omega = 0.06\) a.u.), intensity \(I = 4.3 \times 10^{13}\) W/cm\(^2\). A laser shape

\[
A(t) = \begin{cases} 
\frac{A_0}{\pi} (1 - \cos(\frac{\omega t}{N_e})) \sin(\omega t) & \text{if } 0 \leq t \leq 2\pi N_e/\omega \\
0 & \text{if } t > 2\pi N_e/\omega
\end{cases}
\]

(22)

for the vector potential should lead to an electric field similar to the one employed in the experiment with zero carrier envelope phase.

In Fig. 9 (a) the experimental photoelectron probability \(P(E, \theta)\) is plotted in logarithmic scale as a function of the energy and the angle with respect to the laser polarization in the laboratory frame. Electrons are mainly emitted at small angles, and, due to the short nature of the pulse electron emission is asymmetric along the laser polarization axis (at angles close to 0° and 180°).

We performed TDDFT calculations for different angles \(\theta_L\) between the molecular axis and the laser polarization. The molecular geometry was set at the experimental equilibrium interatomic distance \(R_0 = 2.074\) a.u. The Kohn-Sham wavefunctions were expanded in real space with spacing \(\Delta r = 0.38\) a.u. in a simulation box of \(R_A = 35\) a.u. The photoelectron spectra were calculated with FMM having \(R_C = 25\) a.u., and padding factors \(P = 1\), and \(P_N = 4\).

In Fig. 10 the logarithmic ionization probability \(P_{\theta_L}(E, \theta)\) is plotted as a function of energy \(E\) and angle \(\theta\) measured from the laser polarization axis for different values of \(\theta_L\). As the molecular orientation decreases from 90° \(\leq \theta_L \leq 30°\) we observe an increasing suppression of the emission together with a shift of the maximum that moves away from the laser polarization axis. For \(\theta_L = 0°\) the emission is highly enhanced for all angles and peaked along the laser direction. The signature of multi-center emission interference has been predicted to be particularly marked when the laser polarization is perpendicular to the molecular axis\textsuperscript{51,22} (i.e. \(\theta_L = 90°\)). However, the lowest point in energy of such a pattern is predicted for \(\theta = 90°\) and \(E = \pi^2/2R_0^2 \approx 31\) eV, way above the energy window of observable photoelectrons produced by our laser. A stronger and longer laser pulse would be required to extend the rescattering plateau toward higher energies and therefore to reveal the pattern\textsuperscript{23}

In order to reproduce the experimental \(\bar{P}(E, \theta)\), an average over all the possible molecular orientations should be performed. Due to the axial symmetry of the molecule we can restrict the average to \(0 \leq \theta_L \leq 90°\) and integrate all the contributions with the proper probability weight\textsuperscript{21}

\[
\bar{P}(E, \theta) \propto \int_0^{90°} d\theta_L \sin \theta_L P_{\theta_L}(E, \theta).
\]

(23)

We evaluate Eq. (23) by discretizing the integral in a sum for \(\theta_L = 0°, 30°, 60°, 90°\), and display the result in Fig. 10 (b). Even in this crude approximation, and without taking into account focal averaging, the agreement with the experiment is satisfactory and compares favorably to the molecular strong field approximation shown in Fig. 10 (c). The agreement deteriorates for low energies where the importance of the Coulomb tail is enhanced as it is not fully accounted due to limited dimensions of the simulation box. As a matter of fact the agreement greatly increases for higher energies.
the Fourier transforms of the initial state wavefunctions $\tilde{\Psi}_i$. In this approximation the square root of the magnitude can be evaluated by Fermi’s golden rule as

$$\sqrt{P_H(p)} \propto |A_0 \cdot p| \times |\tilde{\Psi}_H(p)|, \quad (26)$$

where $\tilde{\Psi}_H$ corrects for the dependence on the orbital orientation.

In this section we deal with UV ($\omega = 0.78$ a.u.) angular resolved photoemission triggered by weak lasers. When the external field is weak, non-linear effects can be discarded and first order perturbation theory can be applied. In this situation, the momentum resolved PES can be evaluated by Fermi’s golden rule as

$$P(p) \propto \sum_i |\langle \Psi_f | A_0 \cdot \hat{p} | \Psi_i \rangle|^2 \delta(E_f - E_i - \omega), \quad (24)$$

where $|\Psi_i\rangle$ ($|\Psi_f\rangle$) is the initial (final) many-body wavefunction of the system and $A_0$ is the laser polarization axis. The difficulty in evaluating Eq. (24) lies in the proper treatment of the final state, which in principle belongs to the continuum of the same Hamiltonian of $|\Psi_i\rangle$. In the simplest approach, it is approximated by a plane wave (PW). In this approximation the square root of the momentum-resolved PES is proportional to the sum of the Fourier transforms of the initial state wavefunctions $\tilde{\Psi}_i(p)$ corrected by a geometrical factor $|A_0 \cdot \hat{p}|$

$$\sqrt{P(p)} \propto \sum_i |A_0 \cdot \hat{p}| \times |\tilde{\Psi}_i(p)|. \quad (25)$$

If photoemission peaks are well resolved in momentum, individual initial states can be selectively measured. In this case a correspondence between momentum-resolved PES and electronic states in reciprocal space can be established. The range of applicability of the PW approximation has been discussed in the literature. It has been postulated that Eq. (25) should be valid for (i) $\pi$-conjugated planar molecules, (ii) constituted by light atoms (H, C, N, O) and for (iii) photoelectrons emerging with momentum $\hat{p}$ almost parallel to the polarization axis.

Here we restrict ourselves to photoemission from the highest occupied molecular orbital (HOMO). In this case Eq. (25) becomes

$$\sqrt{P_H(p)} \propto |A_0 \cdot \hat{p}| \times |\tilde{\Psi}_H(p)|, \quad (26)$$

the subscript $H$ indicating HOMO-related quantities. We compare ab-initio TDDFT and PW PADs evaluated at fixed momentum $|p_H| = \sqrt{2E_H}$ with $E_H = \omega - E_B$ being the kinetic energy of photoelectrons emitted from the HOMO and $E_B$ its binding energy.

TDDFT numerical calculations are carried out on a grid with spacing $\Delta r = 0.28$ a.u. for benzene and $\Delta r = 0.38$ a.u. for CO, in a simulation box of $R_A = 36 a.u.$ Photoelectron spectra are calculated using MM potential的能量$W$ with $R_C = 20$ a.u. and padding factors $P = 1, P_{\pi} = 8$. A 40 cycles pulse with 8 cycle ramp at the He-(I)-frequency $\omega = 0.78$ a.u. and intensity $I = 1 \times 10^8$ W/cm$^2$ is employed.

We begin presenting the case of benzene since it constitutes the smallest molecule meeting all the conditions for Eq. (26) to be valid. Results for molecules oriented according to Fig. (11) (a), evaluated at $E_H = 0.363$ a.u., and two different laser polarizations $A_0 = \hat{a}_1, \hat{a}_2$ with $\hat{a}_1 = (1, 0, 0), \hat{a}_2 = 1/\sqrt{3} \times (1, 1, 1)$, are shown in Fig. (12). In the case where the laser is polarized along the $x$ axis [see Fig. (12 (b)), PAD presents a four lobes symmetry separated by three horizontal and two vertical nodal lines. This structure is reminiscent of the HOMO $\pi$-symmetry with the nodal line at $\theta = 90^\circ$ corresponding to the nodes of the orbital on the $x$-$y$ plane. Information on the orientation of the molecular plane could then be inferred from the inspection of this nodal line in the PAD. A similar feature can be observed also in the case of an off-plane polarization as shown in Fig. (12 (d)). In this case, however, the laser can also excite $\sigma$-orbitals and the nodal line at $\theta = 90^\circ$ is partially washed out. The other nodal lines can be understood in term of zeros of the polarization factor $|A_0 \cdot \hat{p}|$ and are thus purely geometrical.

FIG. 10. (Color online) Photoelectron angle and energy resolved probability $P_{\theta_i}(E, \theta)$ (log scale) for aligned $N_2$ molecules and different laser polarization directions $\theta_L$: (a) $\theta_L = 90^\circ$, (b) $\theta_L = 60^\circ$, (c) $\theta_L = 30^\circ$, (d) $\theta_L = 0^\circ$. Here $\theta_L$ is the angle between the laser polarization direction and the molecular axis. Laser parameters are the same as in Fig. [9].

FIG. 11. Photoemission geometries for oriented (a) benzene and (b) CO molecules.
A PW approximation of the photoelectron distribution given by Eq. (26) qualitatively reproduce the ab-initio results as shown in Fig. 12 (a) and (d). According to condition (iii) a quantitative agreement is reached only for directions parallel to the polarization axis.

A different behavior is expected in the case of CO. Photoelectrons with kinetic energy of $E_H = 0.261$ a.u. are show in Fig. 13. In this case, condition (i) (i.e. $\pi$-conjugated molecule) is not fulfilled and a worse agreement is expected. The geometry of the photoemission process is indicated in Fig. 11 (a).

IV. CONCLUSIONS

In this work we studied the problem of photoemission in finite systems with TDDFT. We presented a formal derivation of a photoelectron density functional from a phase-space approach to photoemission. Such a functional can be directly applied to other theories based on a single Slater determinant and the derivation could serve as a base for extensions to more refined models.

We proposed a mixed real- and momentum-space evolution scheme based on geometrical splitting. In its com-
plete form it allows particles to seamlessly pass back and forth from a real-space description to a momentum-space description. The ordinary splitting scheme turns out to be a special case of this more general method. Furthermore, we illustrated applications of the method on four physical systems: hydrogen, molecular nitrogen, carbon monoxide and benzene.

For hydrogen we presented a comparison of the different methods. We studied ATI peak formation in a one-dimensional model and ATI angular distributions for a three-dimensional case. The results turned out to be in good agreement with the literature. From the comparison, we derived a prescription to choose the best method based on a classification of the electron dynamics induced by the external field.

We investigated angular-resolved photoemission for randomly oriented N$_2$ molecules in a short intense IR laser pulse. We illustrated the results for four different molecular orientations with respect to the laser polarization. Owing to the symmetry of the problem we were able to combine the results to account for the random orientation. The spectrum for randomly oriented molecules is in good agreement with experimental measurements and is much better than the widely used strong field approximations (with one active electron).

We also studied UV angular resolved photoelectron spectra for oriented carbon monoxide and benzene molecules. We presented numerical calculations for two different directions of the laser polarization and compared with the plane-wave approximation. We found that the plane-wave approximation provides a good description for benzene while failing for CO. Furthermore, we found evidence that the photoelectron angular distribution carries important information on molecular orientation.

The successful implementation of photoelectron density functional presented in this Article paves the road for interesting applications to many different systems for a wide range of laser parameters. To name a few, TDDFT PAD could provide a theoretical tool superior to the plane-wave and the independent atomic center approximations to retrieve molecular adsorption orientation information from experiments. Atto-second pump probe experiments could be simulated ab-initio accounting for many-body effects but with great computational advantage with respect to full many-body methods and better physical description than SAE pictures.

V. ACKNOWLEDGMENTS

Special thanks to Lorenzo Stella for many stimulating discussions and suggestions. We also wish to acknowledge useful discussions and comments from Stefan Kurth, Ilya Tokatly, Matteo Gatti and Franck Lépine.

Financial support was provided by Spanish (FIS2011-65702-C02-01 and PIB2010US-00652), ACI-Promociona (ACI2009-1036), Grupos Consolidados UPV/EHU del Gobierno Vasco (IT-319-07), and the European Research Council Advanced Grant DYNamo (ERC-2010-AdG - Proposal No. 267374). Computational time was granted by i2basque and BSC Red Española de Supercomputación. MALM acknowledges support from the French ANR (ANR-08-CEXC8-008-01).

Appendix A: Overlap integrals

In this section we describe the details of the inclusion of the Kohn-Sham one-body density matrix $[9]$ into Eq. $[4]$. The momentum-resolved photoelectron probability is the sum over all the occupied orbitals of four overlap integrals $\gamma$

$$P(p) = \sum_{i=1}^{\text{occ.}} \gamma_{A,A,i}(p) + \gamma_{A,B,i}(p) + \gamma_{B,A,i}(p) + \gamma_{B,B,i}(p).$$  \hspace{1cm} (A1)

In order to simplify the notation we drop the orbital index $i$ in the overlap integrals and indicate with $v = vv$ the vector $v$ of modulus $v$ and direction $\hat{v}$. In addition, we will consider the simple case where the boundary surface between region A and B is a $d$-dimensional sphere of radius $R_A$.

We start by considering the mixed overlap

$$\gamma_{AB}(p) = \int_B \text{d}R \int \frac{\text{d}s}{(2\pi)^{\frac{d}{2}}} e^{i(p \cdot s)} \psi_A(R + \frac{s}{2}) \psi_B(R - \frac{s}{2})$$  \hspace{1cm} (A2)

where the integration in $B$ is for $R > R_A$ [cf. Fig. I (a)]. It is convenient to work in the coordinates $v = 2R$ and $r = R + s/2$, where the integral takes the form

$$\int_{v>2R_A} \text{d}v \int \frac{\text{d}r}{(2\pi)^{\frac{d}{2}}} e^{i(p \cdot (2r - v))} \psi_A(r) \psi_B^*(v - r).$$  \hspace{1cm} (A3)

We substitute $\psi_B^*$ with its Fourier integral representation

$$\psi_B^*(u - r) = \int \frac{\text{d}k}{(2\pi)^{\frac{d}{2}}} e^{i (u - r) \cdot k} \tilde{\psi}_B(k)$$  \hspace{1cm} (A4)

and after few simple steps we obtain

$$\int \frac{\text{d}k}{(2\pi)^{\frac{d}{2}}} \tilde{\psi}_A(-2p - k) \tilde{\psi}_B^*(k) \int_{v>2R_A} \text{d}v e^{-i(k+p) \cdot v}$$  \hspace{1cm} (A5)

where we successfully disentangled the integration over $\mathbf{v}$ in the second integral. The integral on $v > 2R_A$ can be rewritten as an integral over the whole space, which yields a $d$-dimensional Dirac delta, minus an integral on $v \leq 2R_A$:

$$\int_{v>2R_A} \text{d}v e^{-i(k+p) \cdot v} =$$

$$(2\pi)^d \delta(k + p) - (4\pi R_A)^d \frac{\int_{d/2}^{2R_A} |k + p|^d}{|k + p|^d}$$  \hspace{1cm} (A6)
where $J_n(k)$ is a Bessel function of the first kind. The second term in (A6) is a function centered in $-\mathbf{p}$ and strongly peaked in the region $w = C_d/R_A$ with $C_1 = \pi$, $C_2 \approx 3.83$, $C_3 \approx 4.49$ being the first zeros of the Bessel function $J_{d/2}(k)$. If the region $w$ is small enough we can consider the integrand in $\mathbf{k}$ of (A5) constant and factor out of the integrand $\tilde{\psi}_A(-2\mathbf{p} - \mathbf{k})\psi_B(\mathbf{k})$ evaluated at $\mathbf{k} = -\mathbf{p}$. It is easy to see that

$$ \int \frac{d\mathbf{k}}{(2\pi)^{\frac{d}{2}}} \frac{J_{d/2}(2R_A|\mathbf{k} + \mathbf{p}|)}{|\mathbf{k} + \mathbf{p}|^{\frac{d}{2}}} = 1 $$

(A7)

and, by plugging (A6) in (A5), we have that $\gamma_{A,B}(\mathbf{p}) \approx 0$. By the same reasoning we should expect $\gamma_{B,A}(\mathbf{p}) \approx 0$.

We now turn to the terms containing wavefunction on the same region. In $(\mathbf{v}, \mathbf{r})$ coordinates

$$ \gamma_{A,A}(\mathbf{p}) = \int_{v > 2R_A} d\mathbf{v} \int \frac{dr}{(2\pi)^{\frac{d}{2}}} e^{i\mathbf{p} \cdot (2\mathbf{r} - \mathbf{v})} \psi_A^*(\mathbf{r})\psi_A(\mathbf{v} - \mathbf{r}). $$

(A8)

The product of functions localized in $A$ is not negligible only for $r < R_A$ and $|v - r| < R_A$. Since the integral is carried out for $v > 2R_A$ we can bound $|v - r|$ from below with $R_A|2v - r|/R_A \geq R_A$. This leads to $R_A \leq |v - r| < R_A$ which is satisfied only on the boundary of $A$. Being a set of negligible measure we have $\gamma_{A,A}(\mathbf{p}) = 0$.

Once again, in $(\mathbf{v}, \mathbf{r})$ coordinates

$$ \gamma_{B,B}(\mathbf{p}) = \int_{v > 2R_A} d\mathbf{v} \int \frac{dr}{(2\pi)^{\frac{d}{2}}} e^{i\mathbf{p} \cdot (2\mathbf{r} - \mathbf{v})} \psi_B(\mathbf{r})\psi_B^*(\mathbf{v} - \mathbf{r}) $$

(A9)

can be written as

$$ \gamma_{B,B}(\mathbf{p}) = |\psi_B(\mathbf{p})|^2 - \int_{v < 2R_A} d\mathbf{v} \int \frac{dr}{(2\pi)^{\frac{d}{2}}} e^{i\mathbf{p} \cdot (2\mathbf{r} - \mathbf{v})} \psi_B(\mathbf{r})\psi_B^*(\mathbf{v} - \mathbf{r}) $$

(A10)

where the first integration is in region $A$. Using the localization of $\psi_B$ we see that the integral is non-zero only for $r > R_A$ and $|v - r| > R_A$. As the integration is for $v < 2R_A$ we have that $R_A \leq |v - r| > R_A$ and therefore the double integral in Eq. (A10) is zero.

### Appendix B: Numerical stability and Fourier integrals

A real-space implementation of Eq. (15) involves the evaluation of several Fourier integrals. Such integrals are necessarily substituted by their discrete equivalent, and therefore discrete Fourier transforms (FT) and fast Fourier transforms (FFTs) are called into play. However, evolution methods based on the discrete FT naturally impose periodic boundary conditions. While this is not presenting any particular issue for MM where FT are only used to map real-space wavefunctions to momentum space, it is a source of numerical instability for FMM where the wavefunctions are reintroduced in the simulation box.

The problem is well illustrated by the following one-dimensional example. Imagine a wavepacket freely propagating to an edge of the simulation box with a certain velocity. In MM, when passing through the buffer region, the packet is converted by discrete FT in momentum space and then analytically evolved as a free particle through the edge of the box. In FMM as the wavefunction evolves in momentum spaces it is also transformed back to real space to account for possible charge returns. In this case, instead of just disappearing from one edge, by virtue of the discrete FT periodic boundary conditions, the same wavepacket will appear from the opposite side. It can be easily understood how such an undesirable event can create a feedback leading to an uncontrolled and unphysical build up of the density.

This behavior can be controlled by the use of zero padding. As we know, the Fourier integrals in Eq. (15) involves functions that are, by construction, zero outside the buffer region $C$. We can therefore enlarge the integration domain (having radius $R_A$) by a padding factor $P$, set the integrand to zero in the extended points, obtaining the same result. As a consequence, a wavepacket propagating toward a boundary edge will have to run an enlarged virtual box of radius $R_A = R_A(2P - 1)$ before emerging from the other side. In addition, the smallest momentum represented $\Delta p = 2\pi/P R_A = \Delta p/P$ in the discretized $\psi_B(i, t')$ is reduced by a factor $1/P$ while the highest momentum $p_{max} = \pi/\Delta r$ remains unchanged.

The price to pay here is an increased memory requirement by a factor $P^d$ (where $d$ is the dimension of the simulation box) and is too high for three-dimensional calculations.

A possible way to find a better scaling is offered by the use of non-uniform discrete Fourier Transform and companion fast algorithm NFFT. NFFT allow for the possibility to perform Fourier integrals on unstructured sampling points with, for fixed accuracy, the same arithmetic complexity as FFT. For a detailed description of the algorithm we refer to the literature. The idea is to use the flexibility of NFFT to perform zero padding in a convenient way. Instead of allocating an enlarged box filled with zeros at equally spaced sample positions, we set only one point at $R_A P_N$ (here $P_N$ is the NFFT padding factor) and evaluate the Fourier integral with NFFT. In this way we gain numerical stability for FMM as long as the wavefunctions are contained in a virtual box of $R_A = R_A(2P_N - 1)$ at the price of adding a number of points that scales as $d - 1$ with the dimension of the box. If $N^d$ is the number of grid points in the simulation box, in order to perform zero padding with NFFT one needs to add only $2N^{d-1}$ points.

With this procedure however, not only the smallest momentum $\Delta p$ is reduced by a factor $1/P_N$, but also the highest momentum $p_{max} = (N/2 + 1)\Delta p$ is decreased by the same amount. This turns out to be the limiting factor in the use of NFFT to preserve numerical stability with FMM as the enlargement factor $P_N$ has an upper bound that depends on the escaping electron dy-
In fact, when we evaluate the back-action term Eq. (16b), we assume \( \psi_{B,i}(p, t') \) to be localized in momentum and, in order to preserve numerical consistency, \( P_B \) must be limited by the highest momentum contained in \( \psi_{B,i} \). A combination of ordinary padding and NFFT padding helps to balance the tradeoff between memory occupancy and numerical stability.

Finally, in MM zero padding can be used to increase resolution in momentum.