Stochastic time-dependent DFT with optimally tuned range-separated hybrids: Application to excitonic effects in large phosphorene sheets

Vojtěch Vlček,[1,a)] (ID) Roi Baer,[2,b)] (ID) and Daniel Neuhauser[3,c)]

AFFILIATIONS

[1] Department of Chemistry and Biochemistry, University of California, Santa Barbara, California 93106, USA
[2] Fritz Haber Center for Molecular Dynamics, Institute of Chemistry, The Hebrew University of Jerusalem, Jerusalem 91904, Israel
[3] Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90095, USA

a) Electronic mail: vlcek@ucsb.edu
b) Electronic mail: roi.baer@huji.ac.il
c) Electronic mail: dxn@ucla.edu

ABSTRACT

We develop a stochastic approach to time-dependent density functional theory with optimally tuned range-separated hybrids containing non-local exchange, for calculating optical spectra. The attractive electron-hole interaction, which leads to the formation of excitons, is included through a time-dependent linear-response technique with a nonlocal exchange interaction which is computed very efficiently through a stochastic scheme. The method is inexpensive and scales quadratically with the number of electrons, at almost the same (low) cost of time dependent Kohn-Sham with local functionals. Our results are in excellent agreement with experimental data, and the efficiency of the approach is demonstrated on large finite phosphorene sheets containing up to 1958 valence electrons.

Published under license by AIP Publishing. https://doi.org/10.1063/1.5093707

I. INTRODUCTION

The understanding of optical excitations in materials is essential for developing novel optical and electronic devices.[1,2] However, it is very challenging to calculate optical properties in extended systems. For small molecules, highly correlated approaches are used, including configuration interaction or the equation of motion coupled cluster approach.[3,4] Further, the optical response is typically described by the Bethe-Salpeter equation (BSE).[1,5] These high level techniques are predictive but scale steeply with the number of electrons so they can only be used for relatively small molecules and unit cells.

An alternative to costly many-body techniques is time-dependent density functional theory (TDDFT)[6] that describes excited state energies, geometries, and other properties of small molecules with a relatively moderate computational cost. In principle, TDDFT is exact, but in practice, approximations have to be introduced. The most common is the adiabatic time-dependent Kohn-Sham (TDKS) theory. While TDKS has been applied successfully to a wide range of molecular systems,[1,2] it suffers from many failures, particularly for extended systems, charge-transfer excited states,[7] multiple excitations,[8] and avoided crossings.[9] The most notable problem of TDKS is the inability to capture low-lying excitonic states in bulk.[1,5]

It has been argued that a TDDFT formulation beyond the Kohn-Sham (KS) picture, namely, a TD-GKS (Generalized Kohn-Sham) approach[10] which employs a nonlocal exchange interaction,[11-13] captures the necessary physics to describe excitation in extended systems[14-16] and molecules,[17-19] and accurately predicts the formation of bound excitons. However, the inclusion of the nonlocal exchange in the TDDFT calculations makes them computationally demanding and out of reach for large nanoscale systems. Recently a family of stochastic orbital methods has been developed to describe ground and excited states with the goal of lowering the computational complexity at the cost of introducing a controllable statistical error.[20-25] Specifically relevant to the current

work is a stochastic ground state range-separated hybrid (RSH) DFT method[23] and a stochastic method for the Bethe-Salpeter equation (BSE).[24]

Our first and main aim in this work is to overcome, using stochastic methods, the computational bottleneck in TDDFT with nonlocal exchange. Specifically, we develop a real-time generalized Kohn-Sham method based on a range-separated hybrid (RSH) with a long-range exact exchange operator. The approach has similarities to that proposed for stochastic BSE (where a damped exchange operator was used), but the starting point is different. Here, the starting point is long-range-corrected RSH-DFT, a method which is known to produce excellent charge-transfer states.[26] Therefore, the present long-range exchange TDDFT starts only from a long-range corrected DFT calculation (which is also implemented stochastically), so the resulting approach is self-contained in the DFT/TDDFT framework and does not resort to a separate computation of individual quasiparticle states; this is in contrast to a BSE work where the starting point is a prior calculation of the quasiparticle states which is achievable, e.g., through the stochastic *GW* method.[21,25,27]

The second aim of the paper is then to use the resulting fully *ab initio* stochastic TDDFT method for describing optical excitations in extremely large systems with thousands of electrons. To validate our method, we investigate the optical properties of phosphorene due to its distinctive properties: it exhibits very strong excitonic signatures and its optical spectra are strongly anisotropic.[28–33]

Below, we first (Sec. II) review the basic theory and present our stochastic implementation. In Sec. III, we verify our method by comparing with experiment for PH_3 and the method is then applied to study 2D phosphorene sheets of increasing sizes. Conclusions follow in Sec. IV.

II. THEORY

It is well-known that absorption can be determined, in linear response, from the time-evolution of an induced dipole density (see the Appendix). This time evolution is in principle governed by the time-dependent Schrödinger equation, which is, however, intractable beyond few-electron systems. DFT[34] is a practical approach for recasting the many-electron system as a set of virtual particles interacting via a mean-field exchange-correlation (XC) potential.

The original formulation of the Kohn-Sham (KS)[35] DFT scheme describes the xc interactions by a local potential which is, in principle, nonlocal in time. In practice, it is further approximated, e.g., by (semi)local functionals of the electronic density at a given time (i.e., the functional is adiabatic). As mentioned in the Introduction, this formulation has some notable failures, e.g., missing excitonic effects. An alternative route, which we pursue here, is to employ a GKS scheme[10,11,36] with nonlocal long-range asymptotic behavior as required for correct description of charge transfer and polarizability[37] and for electron-hole bound states.[12,14–16]

We first review below the formulation of DFT and TDDFT with long-range nonlocal exchange, and then detail the stochastic implementation.

A. DFT with long-range nonlocal exchange

The GKS Hamiltonian reads

$$H[n, \rho] \equiv h_0 + v_H[n(\mathbf{r})] + v_C^\gamma[n(\mathbf{r})] + X^\gamma[\rho(\mathbf{r}, \mathbf{r}')], \qquad (1)$$

where h_0 contains the kinetic energy and the electron-nuclear attraction. The density-density repulsion is given by the Hartree potential v_H, and v_C^γ (where γ is defined shortly) is a (semi)local correlation density functional—we use here a local functional form.[38] The nonlocal exchange interaction X^γ is a functional of the density-matrix $\rho(\mathbf{r}, \mathbf{r}')$, where the density is of course $n(\mathbf{r}) = \rho(\mathbf{r}, \mathbf{r})$. (Note that we use different symbols, n and ρ, for the density and density matrix since later we calculate the two separately—one deterministically and the other stochastically.)

The form of X^γ derives from a screened Coulomb interaction governed by a single parameter γ.[12,36,39] Specifically, the Coulomb kernel is partitioned as

$$\frac{1}{r} = \frac{\text{erfc}(\gamma r)}{r} + \frac{\text{erf}(\gamma r)}{r}, \qquad (2)$$

where γ is the range separation parameter. The first term dominates at small distances ($r \to 0$) and its contribution to the exchange is approximated by a local density functional.[38–40] The second term in Eq. (2) is active at large distances and gives the nonlocal Fock exchange (X_{nl}^γ). The matrix element of the exchange vector is a direct product of the density matrix and the nonlocal Coulomb interaction (the simple product form is important later in the stochastic formulation)

$$X_{nl}^\gamma(\mathbf{r}, \mathbf{r}') \equiv -v(\mathbf{r}, \mathbf{r}')\rho(\mathbf{r}, \mathbf{r}'). \qquad (3)$$

Here, the long-range interaction is $v(\mathbf{r}, \mathbf{r}') = \text{erf}(\gamma|\mathbf{r} - \mathbf{r}'|)/|\mathbf{r} - \mathbf{r}'|$, and the density matrix is determined from the eigenstates $\rho(\mathbf{r}, \mathbf{r}') = \sum_i f_i \phi_i(\mathbf{r})\phi_i(\mathbf{r}')$, where i is a state and spin index and f_i are the occupation factors, while the density is $n(\mathbf{r}) = \rho(\mathbf{r}, \mathbf{r}) = \sum_i f_i |\phi_i(\mathbf{r})|^2$ (In the following, we do not denote spin explicitly). Thus, deterministically, the nonlocal exchange term acts on a general function ψ as

$$\langle \mathbf{r}|X_{nl}^\gamma|\psi\rangle = -\sum_i f_i \phi_i(\mathbf{r}) \int v(\mathbf{r}, \mathbf{r}')\phi_i(\mathbf{r}')\psi(\mathbf{r}')d\mathbf{r}'. \qquad (4)$$

In the first (DFT) stage, the occupied eigenstates $\phi_i(\mathbf{r})$ of Eq. (1) are calculated self-consistently, ensuring $H\phi_i = \varepsilon_i\phi_i$, where the density and density matrix are functions of the eigenstates. The value of γ is found by enforcing the IP theorem that requires that the HOMO energy equals the ionization energy. This optimal tuning leads to good IPs and fundamental band gaps in finite systems.[41–43]

The method's success stems from the combination of (semi)local functionals that capture correlation effects well at short distances with the nonlocal X_{nl}^γ that guarantees for finite systems the asymptotically correct $1/r$ behavior of the exchange potential which is crucial for proper inclusion of the attractive electron-hole interaction.[1,37]

B. TDDFT with long-range nonlocal exchange

It is well-known (the Appendix) that the absorption spectrum is obtained from a linear-response propagation of the density.

Specifically, for polarized excitation along a unit vector $\hat{\mathbf{e}}$, we apply a small perturbation $\delta v(\mathbf{r}, t) = (\mathbf{r} \cdot \hat{\mathbf{e}}) \delta(t) \Delta$ where Δ is a small constant (typically between 10^{-3} and 10^{-5} a.u.). Then, the system evolves under the time-dependent GKS equation (using $\hbar = 1$)

$$i \big| \dot{\phi}_i(t) \big\rangle = [H[n(t), \rho(t)] + \delta v(\mathbf{r}, t)] |\phi_i(t)\rangle. \qquad (5)$$

To simplify the notation, we usually do not denote the dependence of the density and density matrix (and therefore of the time-dependent Hamiltonian) on the excitation strength Δ.

The Hamiltonian $H(t)$ is time dependent as it explicitly depends on the propagated eigenstates $\phi_i(\mathbf{r}, t)$, the time-dependent charge density $n(\mathbf{r}, t) = \sum_i f_i |\phi(\mathbf{r}', t)|^2$, and the charge density matrix, $\rho(\mathbf{r}, \mathbf{r}', t) = \sum_i f_i \phi(\mathbf{r}, t) \phi_i^*(\mathbf{r}', t)$. The dipole moment along the excitation direction is then calculated from the density, $\mu(t) = \int (\mathbf{r} \cdot \hat{\mathbf{e}}) n(\mathbf{r}, t) d\mathbf{r}$, and the absorption spectrum is calculated by Fourier transforming the dipole moment $\mu(t)$ (the Appendix).

In principle, the exchange-correlation term in the Hamiltonian should account for memory effects, but since its form is unknown, we resort to the adiabatic approximation and construct the xc terms directly from $n(\mathbf{r}, t)$ and $\rho(\mathbf{r}, \mathbf{r}', t)$. Thus, the difference from Kohn-Sham type adiabatic TDDFT is only in the exchange kernel.

The application of the nonlocal exchange as presented in Eq. (3) is computationally demanding, due to the integral over the density matrix. Practical calculations are therefore limited to systems with a low number of states.

C. Stochastic DFT with nonlocal exchange

Next, we review our implementation[23] of the DFT equations with a stochastic representation of the nonlocal exchange operator. This is followed by the implementation of stochastic TDDFT in Sec. II D.

The first step is the DFT ground-state calculation, where we use the stochastic-exchange approach of Ref. 23. This grid-based method is done by two key parts. The first is the representation of the density matrix as an average over stochastic correlation functions. Specifically, we construct stochastic states $\{\eta\}$, each of which is a linear combination of all the occupied eigenstates $\{\phi_i\}$ (cf. Refs. 20, 22, 24, 44, and 45)

$$\eta(\mathbf{r}) = \sum_i \sqrt{f_i} \phi_i(\mathbf{r}) \langle \phi_i | \bar{\eta} \rangle, \qquad (6)$$

where $\bar{\eta}$ is a completely random real vector, e.g., $\bar{\eta}(\mathbf{r}) = \pm (dV)^{-1/2}$, and dV is the grid volume-element. It is straightforward to show that as an operator, the density matrix becomes an average over the separable terms

$$\rho = \{|\eta\rangle\langle\eta|\}_{\bar{\eta}}, \qquad (7)$$

i.e., $\rho(\mathbf{r}, \mathbf{r}') = \{\eta(\mathbf{r}) \eta(\mathbf{r}')\}_{\bar{\eta}}$, where $\{\cdots\}_{\bar{\eta}}$ denotes a statistical average over all random states $\bar{\eta}$. Since the average of $|\eta\rangle\langle\eta|$ yields the density matrix, we can view $\eta(\mathbf{r})$ as a stochastic density amplitude.

In the ground-state DFT stage, we supplement the stochastic representation of the density matrix by a similar stochastic decomposition of the long-range Coulomb interaction [Eq. (7)] using stochastic states ζ

$$\frac{\text{erf}(\gamma|\mathbf{r} - \mathbf{r}'|)}{|\mathbf{r} - \mathbf{r}'|} = \{\zeta(\mathbf{r}) \zeta(\mathbf{r}')\}_{\bar{\theta}} \qquad (8)$$

that are evaluated as

$$\zeta(\mathbf{r}) = \frac{1}{V^{\frac{1}{2}}} \sum_{\mathbf{k}}' \sqrt{v(\gamma, \mathbf{k})} e^{i(\bar{\theta}(\mathbf{k}) + \mathbf{k} \cdot \mathbf{r})}, \qquad (9)$$

where $\bar{\theta}(\mathbf{k})$ is a random phase, and we impose $\bar{\theta}(-\mathbf{k}) = -\bar{\theta}(\mathbf{k})$ to ensure that $\zeta(\mathbf{r})$ are real. Also, V is the total volume. The prime in the summation indicates that the $\mathbf{k} = 0$ term is excluded and is later added analytically. The γ-dependent long-range Coulomb interaction in momentum space is $v(\gamma, \mathbf{k}) = 4\pi e^{-\frac{|\mathbf{k}|^2 \gamma^2}{4}} / |\mathbf{k}|^2$. Note that the average in Eq. (8) is over the random phases $\bar{\theta}(\mathbf{k})$ which determine the stochastic function $\zeta(\mathbf{r})$.

The stochastic decompositions of the density matrix and of the Coulomb potential are then combined to give

$$\langle \mathbf{r} | X_{\text{nl}}^\gamma | \phi \rangle = -\{\xi(\mathbf{r}) \langle \xi | \phi \rangle\}_{\bar{\eta}\theta} \simeq -\frac{1}{N_\xi} \sum_{j=1}^{N_\xi} \xi^j(\mathbf{r}) \langle \xi^j | \phi \rangle, \qquad (10)$$

where the combined exchange-operator stochastic amplitude is simply $\xi(\mathbf{r}) = \eta(\mathbf{r}) \zeta(\mathbf{r})$. The average is done now over N_ξ random states; each sampling (labeled by j) of random $\xi(\mathbf{r})$ amounts to a simultaneously choosing (independently) both the phases $\theta(\mathbf{k})$ and the random vector $\bar{\eta}(\mathbf{r})$, so $\xi^j(\mathbf{r}) = \eta^j(\mathbf{r}) \zeta^j(\mathbf{r})$.

Equation (10) is formally exact if the number of states $N_\xi \to \infty$. For any finite number N_ξ there is a statistical error proportional to $1/\sqrt{N_\xi}$, but since the long-range exchange vector is not numerically large, this error is small even when N_ξ is only a few hundreds. Further details, such as the supplementary use of a deterministic HOMO/LUMO when extracting γ, are given in Ref. 23.

Note that in this present approach, the only operator which is stochastically sampled is the long-range exchange. The density is still sampled deterministically from the eigenstates, $n(\mathbf{r}) = \sum_i f_i |\phi_i(\mathbf{r})|^2$, and the DFT cost is similar to that of traditional deterministic DFT for semilocal functionals. In practice, one can use any usual DFT algorithm to iteratively solve $H\phi_i = \varepsilon_i \phi_i$ for the occupied states, with H constructed from $n(\mathbf{r})$ and from X_{nl}^γ in Eq. (10).

The statistical errors in this mixed approach, where only the exchange is sampled stochastically, are much smaller than in our fully stochastic DFT approach[20,46] where the stochastic orbitals $\eta(\mathbf{r})$ were also used to sample the local density (i.e., where we use $n(\mathbf{r}) = \{|\eta(\mathbf{r})|^2\}$, or more generally $n(\mathbf{r}) = n_0(\mathbf{r}) + \{|\eta(\mathbf{r})|^2\}$, where $n_0(\mathbf{r})$ is a fragment density). In the fully stochastic approach, the eigenstates do not need to be determined[20,44,45] so it formally scales linearly with system size; here, the scaling goal is more modest, just to reduce the cost to that of traditional DFT (and later TDDFT) with only local and semilocal potentials.

D. Stochastic TDDFT with nonlocal exchange

Following the DFT stage with stochastic exchange, we turn to the implementation of stochastic-exchange in TDDFT. Due to the evolution in time, the eigenstates acquire a complex phase and it is thus not possible to employ Eq. (6) which used real stochastic vectors. Instead, complex stochastic vectors are used. Specifically, at each time-step, we represent the density matrix as an average over stochastic vectors, where each one is constructed from the occupied eigenstates

$$\beta(\mathbf{r}, t) = \sum_j e^{i\alpha_j(t)} \sqrt{f_j} \phi_j(\mathbf{r}, t), \qquad (11)$$

where $\alpha_i(t) \in [0, 2\pi]$ is a random phase. Thus, each β is a stochastic vector created using a distinct set of random phases $\{\alpha_i\}$, and a different set of random phases is taken as each time step. Obviously, $\{\beta(\mathbf{r}, t)\beta^*(\mathbf{r}', t)\} = \rho(\mathbf{r}, \mathbf{r}', t)$.

Note that the η and β vectors have a similar meaning; the former is used for the initial time-independent stage and the latter for TDDFT. We use a different symbol to emphasize that the number of such stochastic vectors is different in DFT and TDDFT. Specifically, since each time-step is small, the effect of the stochastic exchange per time step is numerically small, so that it is sufficient to use only a small number (N_β) of vectors in each time step. For that reason, we have not done a stochastic resolution of the Coulomb kernel for the time-dependent exchange, which is formally calculated now as

$$\langle \mathbf{r}|X_{\mathrm{nl}}^\gamma(t)|\phi\rangle = -\frac{1}{N_\beta}\sum_{l=1}^{N_\beta}\beta^l(\mathbf{r}, t)\int v(\mathbf{r}-\mathbf{r}')\beta^{l*}(\mathbf{r}', t)\phi(\mathbf{r}')d\mathbf{r}'. \quad (12)$$

Thus, the cost of evaluating Eq. (12) is only N_β-times more expensive than evaluating the Hartree term.

Since Eq. (12) is evaluated stochastically, $H(t)$ exhibits fluctuations even when there is no perturbation. For a linear response in Δ, we therefore need to propagate two equations, with and without the perturbation,

$$i\left|\dot{\phi}_i^\Delta(t)\right\rangle = \left[H^\Delta(t) + \delta v(\mathbf{r}, t)\right]\left|\phi_i^\Delta(t)\right\rangle, \quad (13)$$

$$i\left|\dot{\phi}_i^{\Delta=0}(t)\right\rangle = H^{\Delta=0}(t)\left|\phi_i^{\Delta=0}(t)\right\rangle, \quad (14)$$

where $\phi_i^\Delta(\mathbf{r}, t=0) = \phi_i^{\Delta=0}(\mathbf{r}, t=0) = \phi_i(\mathbf{r})$. $H^\Delta \equiv H[n^\Delta(t), \rho^\Delta(t)]$ and $H^{\Delta=0} \equiv H^\Delta \equiv H[n^{\Delta=0}(t), \rho^{\Delta=0}(t)]$ have the same functional dependence on the density matrix, but since the time-dependent solutions of Eqs. (13) and (14) are different, we distinguish the Hamiltonians by superscripts. The time evolution of $\{\phi_i^{\Delta=0}\}$ stems purely from the stochastic fluctuations in $H^{\Delta=0}(t)$, as no external perturbing potential is applied. This fluctuation also induces time-dependence in the charge density $n^{\Delta=0}(\mathbf{r}, t)$, which needs to be subtracted when calculating the induced dipole

$$\mu(t) = \frac{1}{\Delta}\int(\mathbf{r}\cdot\hat{\mathbf{e}})\left[n^\Delta(\mathbf{r}, t) - n^{\Delta=0}(\mathbf{r}, t)\right]d\mathbf{r}, \quad (15)$$

from which the frequency-dependent absorption follows.

E. Numerical propagation of TDDFT with stochastic exchange

We use a split operator approach for the numerical propagation of the TDDFT equation with stochastic exchange. As usual, the perturbation is first applied at $t = 0$ (and we again omit below the Δ superscript)

$$\phi_i(\mathbf{r}, t=0^+) = e^{-i(\mathbf{r}\cdot\hat{\mathbf{e}})\Delta}\phi_i(\mathbf{r}), \quad (16)$$

and we then split the propagation of the nonlocal exchange and the remainder of the Hamiltonian

$$|\phi_i(t+dt)\rangle = e^{-iX_{\mathrm{nl}}^\gamma\frac{dt}{2}}e^{-i(h_0+v_H[n(t)]+v_C[n(t)])dt}e^{-iX_{\mathrm{nl}}^\gamma\frac{dt}{2}}|\phi_i(t)\rangle. \quad (17)$$

The short time kinetic+potential propagator [the non-X_{nl}^γ part in Eq. (17)] is itself calculated with a usual split operator evolution

which will not be reviewed here, while $e^{-iX_{\mathrm{nl}}^\gamma\frac{dt}{2}}$ is evaluated extremely simply as

$$e^{-iX_{\mathrm{nl}}^\gamma\frac{dt}{2}}|\phi\rangle \simeq N_\phi\left(1 - iX_{\mathrm{nl}}^\gamma\frac{dt}{2}\right)|\phi\rangle, \quad (18)$$

where N_ϕ is a time-dependent normalization constant, i.e., $N_\phi^{-1} = \|\left(1 - iX_{\mathrm{nl}}^\gamma\frac{dt}{2}\right)|\phi\rangle\|$. Since the normalization is dependent on the initial vector, Eq. (18) is slightly nonlinear, but this is of little practical consequence. The primitive approach of Eq. (18) is sufficient since the time-steps used are generally small, typically $dt = 0.05$ a.u., i.e., around 1 as.

III. RESULTS

For all systems studied here, we first perform a ground-state DFT calculation and obtain the range-separation parameter γ by enforcing the piecewise linearity condition for the total energy; this ensures that the HOMO is the same as the ionization energy.

A. Validation of the method using PH₃

The smallest system studied is a PH_3 molecule. Here, a deterministic DFT calculation was performed using experimental molecular structure,[47] and the valence electronic states were computed with Troullier-Martins pseudopotentials.[48] Note that stochastic sampling was not used for the ground state calculation, but it was applied for TDDFT as detailed below. The total energy and the eigenvalues were converged to 5 meV with a real space grid of 64 × 64 × 64 points and a 0.4 a_0 grid spacing. Note that small molecular systems require, in general, a large range separation parameter and converge slower with the grid size and spacing compared to large systems. Through the tuning procedure,[23,41,43] we found that $\gamma = 0.37\ a_0^{-1}$; the resulting ionization potential (i.e., the negative of the HOMO energy) is 10.4 eV, in excellent agreement with experiment (10.6 eV—Ref. 49).

The LUMO (obtained with the same range separation parameter) is barely bound, by slightly less than 0.1 eV, but experimentally PH_3 does not form a stable anion so the LUMO energy should be non-negative.

Using the optimally tuned BNL functional, the optical cross-section $\sigma(\omega)$ was obtained [see Eq. (A6)] by deterministic and stochastic real time propagations, and the results are shown in Fig. 1. The TDDFT equations were propagated for a total time of 24 fs, which provides a spectral resolution of ~170 meV. The computed absorption crosssection has a first peak at $E_1 = 7.1$ eV, in excellent agreement with the first experimental peak at 7.0 eV.

The exciton binding energy is defined as

$$E_b = E_g - E_1, \quad (19)$$

where E_g is the fundamental band gap taken as the difference between HOMO and LUMO energies, i.e., $E_g = 10.3$ eV. The predicted PH_3 binding energy is, thus, $E_b = 3.2$ eV, close to the experimental value of 3.4 eV.[50]

The overall absorption maximum is at 11.5 eV, which is in good agreement with experiment (12.0 eV), though the latter exhibits large peak widths. At higher frequencies, the TDDFT spectrum has multiple local maxima (e.g., at 16.5 and 20.4 eV) that in the experiment only appear as shoulders. This is because experimental
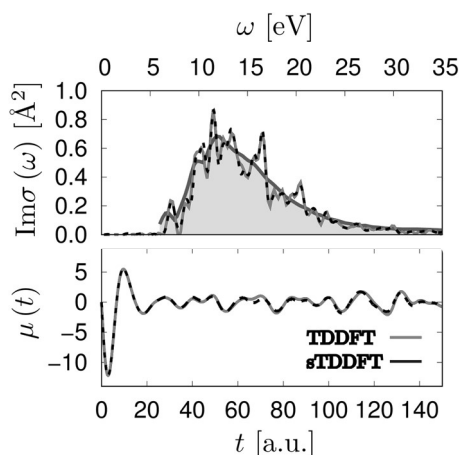
FIG. 1. Top: Optical absorption of a PH_3 molecule from stochastic and deterministic TDDFT simulations (full and dashed lines, respectively). The experimental spectrum[51] is shown by a blue shaded area. Bottom: An initial segment of the time propagation showing the evolution of the induced dipole $\mu(t)$.

measurements cannot be resolved at energies higher than the ionization threshold (10.4 eV).

The stochastic decomposition of the time-dependent exchange [sTDDFT—Eq. (12)] reproduces the deterministic results already for $N_\beta = 2$ (note that this is half the number used in deterministic exchange, which involves four valence states). We also checked that the spectrum does not change when the perturbation is varied in the range $\Delta = 10^{-4}$–10^{-3} a.u., as expected in linear response; this was also checked for the phosphorene sheets, which will be discussed next.

B. Phosphorene sheets: DFT

The major advantage of stochastic approaches is their applicability to large systems. We demonstrate this feature now on a set of 2D phosphorene sheets of increasing sizes, derived from a black phosphorus crystal structure.[52] The sheets were passivated with H atoms on the rims; each P atom that would have been bound to two H atoms was removed, resulting in a compact sheet geometry. Note that the two in-plane directions in the phosphorene sheet are traditionally labeled as armchair and zig-zag. With a kinetic energy cutoff of 26 E_h and a real space grid with $208 \times 136 \times 40$ points and a 0.6 a_0 spacing, the Kohn-Sham eigenvalues were converged to 10 meV.

The smallest sheet is 0.6×0.8 nm^2 and has 112 valence electrons. For its ground state DFT calculations, we employed both the deterministic and stochastic formulation of the exchange operator X_{nl}^γ. The optimally tuned range-separation parameter for this sheet is $\gamma = 0.10 \, a_0^{-1}$, and the stochastic eigenvalues converge slowly with the number of stochastic states, so $N_\xi \sim 1600$ is required to yield a statistical error of <0.05 eV.

In addition to the small sheet, we considered two larger sheets, 1.3×2.1 nm and 3.1×4.3 nm (labeled "medium" and "big"), with 478 and 1958 valence electrons, respectively. For these larger sheets, the exchange operator was calculated purely stochastically as the

deterministic calculation would have been very expensive. The range separation parameter gradually decreases with system size as in other 1D and 3D systems,[41,53,54] so $\gamma = 0.09 a_0^{-1}$ in the medium-size sheet and $0.05 a_0^{-1}$ in the largest one. As the range separation parameter decreases with system size, the long-range exchange operator X_{nl}^γ is numerically smaller and its stochastic representation has therefore a small absolute statistical error. Hence, the largest system requires a smaller value of N_ξ in the ground state calculations. Namely, the eigenvalues are converged to <50 meV with $N_\xi = 1600$ stochastic states for the medium-size sheet and only $N_\xi = 400$ for the largest one.

The stochastic-exchange DFT yields fundamental gaps E_g that decrease with system size: $E_g = 3.9, 3.1$, and 1.7 eV for the three sheets respectively. The large-sheet result is very similar to the HSE hybrid functional prediction of 1.5 eV.[55] Comparing to with previous periodic $G_0 W_0$ calculations (with a starting point based on a semilocal density functional), we find that the largest sheet is in rough but not perfect agreement with the 2D periodic $G_0 W_0$ fundamental gap of 2.08 eV.[33,56] The difference from the GW result could be due to the scalar (i.e., nondirectional) nature of the range separated parameter, which ignores the difference between the effective interactions in the in- and out-of plane directions, and perhaps also due to the approximate nature of the $G_0 W_0$ itself.

C. Phosphorene sheets: TDDFT results

We next discuss the stability of TDDFT simulations for the sheets and the resulting optical spectra. Since the wavefunction is only incremented gradually, by $dt/2 = 0.025$ a.u., the statistical fluctuation introduced by each stochastic decomposition of the exchange operator is significantly smaller than for the ground state calculation. Hence, as mentioned, a small N_β is sufficient for the time-dependent calculation, so the short-time results are fairly accurate already for $N_\beta = 1$–4. However, N_β influences the total time of the simulations since due to statistical fluctuations the propagation eventually becomes unstable. We verified that this instability is not influenced by the time step, grid size, and kinetic-energy cut-off.

We noticed the instability phenomena already in our original stochastic TDDFT approach[22] where, unlike here, we propagated only a few stochastic combination of eigenstates [i.e., several $\beta(\mathbf{r}, t)$], and constructed from them the density as $n(\mathbf{r}, t) \simeq N_\beta^{-1} \sum_\beta |\beta(\mathbf{r}, t)|^2$. That approach is extremely efficient for short time simulations (where the plasmon response of systems with thousands of electrons is accurately modeled by circa 10 propagated states), but is limited to short times since the propagation eventually becomes unstable. Here, since all occupied eigenstates are propagated and the density is constructed from all of them, the propagation is fairly stable for longer times.

Specifically, for the small and medium sheets, with 112 and 478 valence electrons, the propagation was carried up to 1000 a.u. (~24 fs) without stability issues, using $N_\beta = 2$. This is in line with our previous simulations (Ref. 24) which used a damped exchange (reduced by 80%) in 3D and were stable with $N_\beta = 1$. However, for the large sheet (with 1958 electrons), the time evolution became numerically unstable after 260 a.u. (~7 fs) even with $N_\beta = 6$. The instability is for two reasons: first, the rapid oscillations of the density in the direction perpendicular to the phosphorene sheet. Unlike 3D systems, the response here is highly anisotropic and this appears
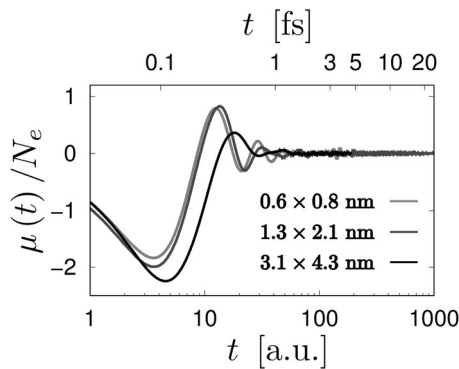
FIG. 2. Induced dipole per electron, $\mu(t)/N_e$, along the armchair direction of phosphorene, plotted for three phosphorene sheets with different lateral dimensions. The largest system (black) was propagated only to 210 a.u., after which the stochastic fluctuations dominate the signal. The periods of the induced dipole oscillations grow with system size. A logarithmic time axis is used; the oscillations are mostly nonstochastic and due to the logarithmic axis.

to enhance the stochastic noise in the time propagation. Hence, for large 2D sheets the value of N_β needs to be increased. Further, the form of Eq. (12) is oscillatory even for $\Delta = 0$; in future publications, we would use a less oscillatory form analogous to that in Ref. 24 and, in addition, would use a fully separable form of the TDDFT calculations, analogous to Eq. (10) for DFT with exchange.

Next, we turn to discuss the individual results. The time evolution of the induced dipole for the three phosphorene sheets is shown in Fig. 2. The absorption cross sections per-electron are shown in Fig. 3 for the zig-zag and armchair directions. The spectra are strongly anisotropic, in agreement with experimental data and GW/BSE calculations on 2D periodic sheets.[28,29,31,33] With rising system size, there are diminished quantum confinement effects so the fundamental band gap decreases and the absorption maximum, therefore, gradually shifts to lower energies. As the number of valence electrons increases, the spectra also smoothen.

On the right panel of Fig. 3, we zoom on the absorption spectrum below the ionization potential (which only slowly decreases

with system size, from 5.9 to 5.4 eV). The absorption cross section decreases rapidly at lower frequencies, but several local maxima are evident. Some of these local maxima are below the band gap energy (E_g) and therefore correspond to bound electron-hole pairs. These excitonic peaks appear only for the armchair direction (due to the strong anisotropy of the optical response); this feature was seen in previous calculations for periodic phosphorene and was also seen experimentally.[28,30,33]

Excitonic peaks are usually sharp and have a high intensity, indicating long-lived quasiparticle states. The maxima in Fig. 3 are, however, broadened due to the finite simulation time (24 fs for the two small systems and 7 fs for the largest one). For the small sheet, the excitonic peaks are well-separated but have relatively low intensity. The position of the first absorption peak maximum (E_1) changes with increasing system size from 1.8 to 1.6 eV. The latter is in good agreement with the experimental value and GW/BSE estimates, 1.7 and 1.6 eV, respectively,[33,56] for a bulk 2D system.

The exciton binding energy [Eq. (19)] decreases rapidly with system size from 2.1 to 0.1 eV. The strongest excitonic response (the largest amplitude of the E_1 peak) is found in the medium sized system, which also has a high exciton binding energy $E_b = 1.65$ eV. As mentioned, however, the fundamental gap E_g in the stochastic-exchange DFT is underestimated relative to G_0W_0 calculations. Therefore, the exciton binding energy for the largest sheet (0.1 eV) is much lower than predicted by GW/BSE calculations, which give $E_b = 0.48$ eV.[56]

Interestingly, when the phosphorene is encapsulated in dielectric media, the GW/BSE binding energy becomes small, 0.14 eV,[56] comparable to our TDDFT estimates of pristine (nonencapsulated) phosphorenes. The encapsulation causes strong screening above and below the 2D system (i.e., in the out-of-plane direction) leading to a big change in the G_0W_0 gap E_g (from 2.08 to 1.62 eV, similar to our pristine DFT gap), while the position of the first excitonic peak, E_1, remains practically unaffected.[56] The difference between our results and experiment and GW/BSE points to a problem in describing 2D materials with range-separated potentials, which should, in principle, account for the anisotropy of the electron-electron interaction between the in- and out-of-plane directions.
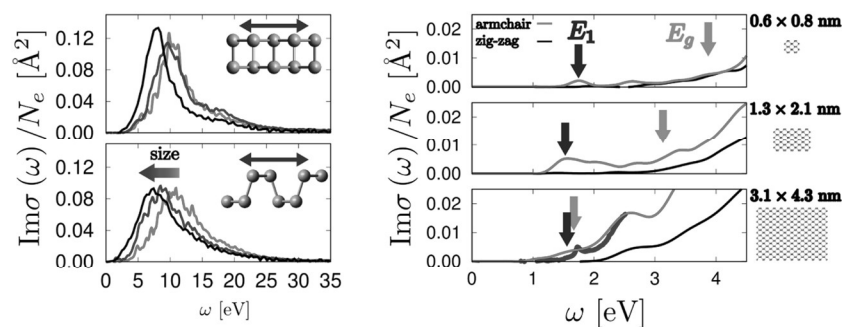

FIG. 3. The left panel shows the theoretical optical spectra of phosphorene sheets of increasing sizes: 0.6×0.8 nm^2 (112 valence electrons), 1.3×2.1 nm^2 (478 valence electrons), and 3.1×4.3 nm^2 (1958 valence electrons) marked by red, blue, and black lines, respectively, for the zig-zag (top) and armchair (bottom) directions. The right panel shows details of the low energy portion of the spectra; the yellow and dark-green vertical lines indicate the positions of the fundamental band gap (E_g) and the first excitonic peak (E_1). Experimental data for a bulk 2D monolayer phosphorene (taken from Ref. 33) are shown in the bottom graph by blue points. Features that lie between E_1 and E_g correspond to multiple excitonic states.

In systems with fairly isotropic electron-electron interaction, which are well approximated by a scalar range-separation parameter, we expect DFT to yield good fundamental band gaps and TDDFT to provide good optical absorption spectra (including excitonic signatures). Hence, for such systems, the exciton binding energies can be, in principle, predicted. However, for highly anisotropic systems stochastic TDDFT with isotropic long-range exchange (and therefore our stochastic version) can only be trusted for the exciton frequency, not the binding energy. In this case, calculations of the exciton binding would require a combination with more accurate methods for the quasiparticle gap, or extended to deal with nonisotropic exchange interactions.

IV. CONCLUSIONS

In summary, we developed an efficient real-time TDDFT approach with stochastic long-range nonlocal exchange. The stochastic treatment decomposes the density matrix in TDDFT to an average over a product of random vectors β in the space spanned by the occupied orbitals. It significantly reduces the computational cost as only a few stochastic states are needed at each time step. Further, the number of stochastic states varies only a little with the system size. Calculations for very large systems thus become feasible.

The resulting TDDFT with long-range nonlocal exchange includes the attractive electron-hole interaction that gives rise to exciton formation. Indeed, our TDDFT yields optical spectra that are in excellent agreement with experiment. For small systems, where deterministic calculations are affordable, the stochastic and deterministic results agree.

We demonstrated that our method is applicable for extremely big systems using a set of phosphorene sheets containing up to ~2000 valence electrons. The largest system was compared to experiments and previous calculations on infinite phosphorene sheets; the analysis confirms that the range-separated hybrid functional successfully predicts optical spectra even with strong excitonic signatures.

APPENDIX: PHOTOABSORPTION CROSS SECTION

Here, we overview for completeness the well-known expression of the photoabsorption cross-section as a Fourier transform of a real-time dipole correlation function.

The absorption cross section, $\sigma(\omega)$, is given in linear response as[1]

$$\sigma(\omega) = \frac{4\pi}{c} \omega \iint \delta\tilde{v}(\mathbf{r}, \omega) \tilde{\chi}(\mathbf{r}, \mathbf{r}', \omega) \delta\tilde{v}(\mathbf{r}', \omega) \, \mathrm{d}\mathbf{r} \, \mathrm{d}\mathbf{r}', \qquad (A1)$$

where tilde is used occasionally to denote quantities in frequency domain, $\delta\tilde{v}(\mathbf{r}, \omega)$ is dynamical external potential, and χ is the electronic reducible polarizability, which is given in the time domain

as

$$\chi(\mathbf{r}, \mathbf{r}', t - t') = \frac{\delta n(\mathbf{r}, t)}{\delta v(\mathbf{r}', t')}, \qquad (A2)$$

where $\delta n(\mathbf{r}, t)$ is the induced charge density at a point $\mathbf{r}$ and time t. The response function is causal, i.e., $t > t'$. For absorption of polarized light along a unit vector $\hat{\mathbf{e}}$, we apply δv as a dipole potential. The cross section is then[1]

$$\sigma(\omega) = \frac{4\pi}{c} \omega \iint (\hat{\mathbf{e}} \cdot \mathbf{r}) \tilde{\chi}(\mathbf{r}, \mathbf{r}', \omega) \cdot (\hat{\mathbf{e}} \cdot \mathbf{r}') \, \mathrm{d}\mathbf{r} \, \mathrm{d}\mathbf{r}'. \qquad (A3)$$

Here, σ is calculated from real-time linear-response. Specifically, the first stage is to apply a dipole external potential perturbation

$$\delta v(\mathbf{r}', t) = (\hat{\mathbf{e}} \cdot \mathbf{r}') \delta(t) \Delta, \qquad (A4)$$

where Δ is the perturbation strength, and an instantaneous perturbation is applied at $t = 0$, allowing to probe the response at all frequencies. The perturbation potential is applied to all occupied eigenstates, which are then propagated in time. The resulting oscillations of the induced charge density [Eq. (A2)] are then used to find the dipole autocorrelation

$$\mu(t) = \frac{1}{\Delta} \int (\hat{\mathbf{e}} \cdot \mathbf{r}) \cdot \delta n(\mathbf{r}, t) \, \mathrm{d}\mathbf{r}. \qquad (A5)$$

The absorption cross section is finally a Fourier transform of the dipole auto correlation

$$\sigma(\omega) = \frac{4\pi\omega}{c} \int_0^\infty \mu(t) e^{i\omega t} \, \mathrm{d}t. \qquad (A6)$$

REFERENCES

[1] G. Onida, L. Reining, and A. Rubio, Rev. Mod. Phys. **74**, 601 (2002).

[2] M. A. Marques and E. K. Gross, Annu. Rev. Phys. Chem. **55**, 427 (2004).

[3] R. J. Bartlett and M. Musiał, Rev. Mod. Phys. **79**, 291 (2007).

[4] T. Helgaker, P. Jorgensen, and J. Olsen, *Molecular Electronic-Structure Theory* (John Wiley & Sons, 2014).

[5] R. M. Martin, L. Reining, and D. M. Ceperley, *Interacting Electrons* (Cambridge University Press, 2016).

[6] E. Runge and E. K. U. Gross, Phys. Rev. Lett. **52**, 997 (1984).

[7] A. Dreuw, J. L. Weisman, and M. Head-Gordon, J. Chem. Phys. **119**, 2943 (2003).

[8] N. T. Maitra, F. Zhang, R. J. Cave, and K. Burke, J. Chem. Phys. **120**, 5932 (2004).

[9] B. G. Levine, C. Ko, J. Quenneville, and T. J. Martínez, Mol. Phys. **104**, 1039 (2006).

[10] R. Baer and L. Kronik, Eur. Phys. J. B **91**, 170 (2018).

[11] A. Seidl, A. Görling, P. Vogl, J. A. Majewski, and M. Levy, Phys. Rev. B **53**, 3764 (1996).

[12] R. Baer and D. Neuhauser, Phys. Rev. Lett. **94**, 043002 (2005).

[13] S. Kümmel and L. Kronik, Rev. Mod. Phys. **80**, 3 (2008).

[14] Z. h. Yang, F. Sottile, and C. A. Ullrich, Phys. Rev. B **92**, 035202 (2015).

[15] S. Refaely-Abramson, M. Jain, S. Sharifzadeh, J. B. Neaton, and L. Kronik, Phys. Rev. B **92**, 081204 (2015).

[16] N. P. Brawand, M. Vörös, M. Govoni, and G. Galli, Phys. Rev. X **6**, 041002 (2016).

[17] K. Lopata and N. Govind, J. Chem. Theory Comput. **7**, 1344 (2011).

[18] A. Sissay, P. Abanador, F. Mauger, M. Gaarde, K. J. Schafer, and K. Lopata, J. Chem. Phys. **145**, 094105 (2016).

[19] A. Bruner, S. Hernandez, F. Mauger, P. M. Abanador, D. J. LaMaster, M. B. Gaarde, K. J. Schafer, and K. Lopata, J. Phys. Chem. Lett. **8**, 3991 (2017).

[20] R. Baer, D. Neuhauser, and E. Rabani, Phys. Rev. Lett. **111**, 106402 (2013).

[21] D. Neuhauser, Y. Gao, C. Arntsen, C. Karshenas, E. Rabani, and R. Baer, Phys. Rev. Lett. **113**, 076402 (2014).

[22] Y. Gao, D. Neuhauser, R. Baer, and E. Rabani, J. Chem. Phys. **142**, 034106 (2015).

[23] D. Neuhauser, E. Rabani, Y. Cytter, and R. Baer, J. Phys. Chem. A **120**, 3071 (2015).

[24] E. Rabani, R. Baer, and D. Neuhauser, Phys. Rev. B **91**, 235302 (2015).

[25] V. Vlček, E. Rabani, D. Neuhauser, and R. Baer, J. Chem. Theory Comput. **13**, 4997 (2017).

[26] T. Stein, L. Kronik, and R. Baer, J. Am. Chem. Soc. **131**, 2818 (2009).

[27] V. Vlček, W. Li, R. Baer, E. Rabani, and D. Neuhauser, Phys. Rev. B **98**, 075107 (2018).

[28] H. Liu, A. T. Neal, Z. Zhu, Z. Luo, X. Xu, D. Tománek, and P. D. Ye, ACS Nano **8**, 4033 (2014).

[29] F. Xia, H. Wang, and Y. Jia, Nat. Commun. **5**, 4458 (2014).

[30] V. Tran, R. Soklaski, Y. Liang, and L. Yang, Phys. Rev. B **89**, 235319 (2014).

[31] X. Wang, A. M. Jones, K. L. Seyler, V. Tran, Y. Jia, H. Zhao, H. Wang, L. Yang, X. Xu, and F. Xia, Nat. Nanotechnol. **10**, 517 (2015).

[32] E. Prada, J. V. Alvarez, K. L. Narasimha-Acharya, F. J. Bailen, and J. J. Palacios, Phys. Rev. B **91**, 245421 (2015).

[33] L. Li, J. Kim, C. Jin, G. J. Ye, D. Y. Qiu, H. Felipe, Z. Shi, L. Chen, Z. Zhang, F. Yang et al., Nat. Nanotechnol. **12**, 21 (2017).

[34] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[35] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[36] A. Savin and H.-J. Flad, Int. J. Quantum Chem. **56**, 327 (1995).

[37] P. Ghosez, X. Gonze, and R. Godby, Phys. Rev. B **56**, 12811 (1997).

[38] J. P. Perdew and Y. Wang, Phys. Rev. B **45**, 13244 (1992).

[39] T. Leininger, H. Stoll, H.-J. Werner, and A. Savin, Chem. Phys. Lett. **275**, 151 (1997).

[40] E. Livshits and R. Baer, Phys. Chem. Chem. Phys. **9**, 2932 (2007).

[41] T. Stein, H. Eisenberg, L. Kronik, and R. Baer, Phys. Rev. Lett. **105**, 266802 (2010).

[42] T. Stein, J. Autschbach, N. Govind, L. Kronik, and R. Baer, J. Phys. Chem. Lett. **3**, 3740 (2012).

[43] L. Kronik, T. Stein, S. Refaely-Abramson, and R. Baer, J. Chem. Theory Comput. **8**, 1515 (2012).

[44] R. Baer and D. Neuhauser, J. Chem. Phys. **137**, 051103 (2012).

[45] D. Neuhauser, E. Rabani, and R. Baer, J. Chem. Theory Comput. **9**, 24 (2013).

[46] D. Neuhauser, R. Baer, and E. Rabani, J. Chem. Phys. **141**, 041102 (2014).

[47] G. Herzberg, Molecular Spectra and Molecular Structure. III. Electronic Spectra and Electronic Structure of Polyatomic Molecules (Van Nostrand Reinhold, New York, 1966).

[48] N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991).

[49] A. Cowley, R. Kemp, M. Lattman, and M. McKee, Inorg. Chem. **21**, 85 (1982).

[50] For PH_3, we assume that the fundamental band gap coincides with the ionization potential as the molecule does not bind an extra electron.

[51] E. B. Zarate, G. Cooper, and C. Brion, Chem. Phys. **148**, 277 (1990).

[52] L. Cartz, S. R. Srinivasa, R. J. Riedner, J. D. Jorgensen, and T. G. Worlton, J. Chem. Phys. **71**, 1718 (1979).

[53] T. Körzdörfer, J. S. Sears, C. Sutton, and J.-L. Brédas, J. Chem. Phys. **135**, 204107 (2011).

[54] V. Vlček, H. R. Eisenberg, G. Steinle-Neumann, D. Neuhauser, E. Rabani, and R. Baer, Phys. Rev. Lett. **116**, 186401 (2016).

[55] J. Qiao, X. Kong, Z.-X. Hu, F. Yang, and W. Ji, Nat. Commun. **5**, 4475 (2014).

[56] D. Y. Qiu, F. H. da Jornada, and S. G. Louie, Nano Lett. **17**, 4706 (2017).

[57] J. Towns, T. Cockerill, M. Dahan, I. Foster, K. Gaither, A. Grimshaw, V. Hazlewood, S. Lathrop, D. Lifka, G. D. Peterson et al., Comput. Sci. Eng. **16**, 62 (2014).