Sublinear scaling for time-dependent stochastic density functional theory

Yi Gao and Daniel Neuhauser
Department of Chemistry and Biochemistry, University of California, Los Angeles, CA-90095 USA

Roi Baer
Fritz Haber Center for Molecular Dynamics, Institute of Chemistry, The Hebrew University of Jerusalem, Jerusalem 91904, Israel

Eran Rabani
Department of Chemistry, University of California and Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA

A stochastic approach to time-dependent density functional theory (TDDFT) is developed for computing the absorption cross section and the random phase approximation (RPA) correlation energy. The core idea of the approach involves time-propagation of a small set of stochastic orbitals which are first projected on the occupied space and then propagated in time according to the time-dependent Kohn-Sham equations. The evolving electron density is exactly represented when the number of random orbitals is infinite, but even a small number (≈ 16) of such orbitals is enough to obtain meaningful results for absorption spectrum and the RPA correlation energy per electron. We implement the approach for silicon nanocrystals (NCs) using real-space grids and find that the overall scaling of the algorithm is sublinear with computational time and memory.

I. INTRODUCTION

Time-dependent density functional theory (TDDFT)[1] allows for practical calculations of the time evolution of electronic densities under time-dependent perturbations. In principle TDDFT is an exact theory,[1] but in applications, several assumptions and approximations are typically made.[2–6] For example, for the most common usage of TDDFT, namely the absorption spectrum[7] of molecules and materials[8] one uses the adiabatic approximation local/semi-local exchange-correlation potentials (time-dependent adiabatic local density approximation[9] (TDALDA) or time-dependent adiabatic generalized gradient approximation (TDAGGA)). TDDFT can also be used to compute the ground-state DFT correlation energy within the adiabatic-connection fluctuation-dissipation (ACFD) approach,[10] or for studying strong-field nonpetrurbative dynamics.[11–16]

There are two types of challenges facing the application of TDDFT for large systems. One is the construction of appropriate functionals, as the simplest, local and semi-local adiabatic functionals (TDALDA and TDAGGA) often fail for large systems.[17–22] The second issue is the development of a linear-scaling approach that overcomes not only the quartic ($O\left(N^4\right)$) scaling in the frequency-domain formulation[23] but also the quadratic ($O\left(N^2\right)$) limit achieved when real-time propagation according to the time-dependent Kohn-Sham (TDKS) equations is used.[24–26] This latter scaling is commonly considered the lowest theoretical scaling limit as it does not require full resolution of the TDKS excitation energies. This is important for large systems where the density of excited states is very large and there is no point in resolving of all single-excited states as in small systems.

The present paper addresses the second challenge described above and presents a stochastic formulation of TDDFT (TDsDFT) formally equivalent to the TDKS method but *without* the Kohn-Sham (KS) orbitals. The new method is based on representing the time-dependent density as an average over densities produced by *evolving* projected stochastic orbitals.[27,28] We consider two demonstrations of the TDsDFT within the linear response limit: The first concerns the calculation of the dipole absorption cross section and the second is based on the ACFD approach to calculate the random phase approximation DFT correlation energy .

The paper is organized as follows: Section II first reviews the relation between linear response TDDFT and the generalized susceptibility operator $\hat{\chi}^{\lambda}(t)$. Next, we show how TDsDFT can be used to perform the time consuming computational step in linear response applications, i.e. the action of $\hat{\chi}^{\lambda}(t)$ on a given potential. In Section III we show how the absorption spectrum and the ACFD-RPA correlation energy can be calculated using TDsDFT. We present results for a series of silicon NCs of varying sizes. We also analyze the scaling, accuracy and stability of the proposed TDsDFT. In Section IV we conclude.

II. THEORY

A. The Generalized Susceptibility Function and Time-Dependent Density Functional Theory

Consider a system of N_e electrons interacting via a damped Coulomb potential ($\lambda v_C(|\mathbf{r}-\mathbf{r}'|)$ where $0 \leq \lambda \leq 1$ and $v_C(r) = e^2/4\pi\epsilon_0 r$ in their ground state $|0_\lambda\rangle$ and having a density $n_0(\mathbf{r}) = \langle 0_\lambda | \hat{n}(\mathbf{r}) | 0_\lambda \rangle$. The linear density response of the system at time t ($\delta n^{\lambda}(\mathbf{r},t)$) to

a small external time-dependent potential perturbation ($v(\mathbf{r}',t')$) is described by the following integral:[29]

$$\delta n^{\lambda}(\mathbf{r},t)=\int_0^t dt'\int d\mathbf{r}'\chi^{\lambda}(\mathbf{r},\mathbf{r}',t-t')\,\delta v(\mathbf{r}',t')\,, \quad (1)$$

where $\chi^{\lambda}(\mathbf{r},\mathbf{r}',t)$ is the generalized susceptibility function,[30] which is also given by retarded density-density correlation function of the system:

$$\chi^{\lambda}(\mathbf{r},\mathbf{r}',t)=(i\hbar)^{-1}\theta(t)\left\langle 0_{\lambda}\left|\left[\hat{n}^{\lambda}(\mathbf{r},t),\hat{n}^{\lambda}(\mathbf{r}',0)\right]\right|0_{\lambda}\right\rangle, \quad (2)$$

where $\hat{n}^{\lambda}(\mathbf{r},t)$ is the density operator at position $\mathbf{r}$ and time t. Eq. (2) is also known as the fluctuation-dissipation relation.[31] $\chi^{\lambda}(\mathbf{r},\mathbf{r}',t)$ is used, for example, to compute the linear polarizability and energy absorption of the system under external fields, the dielectric response, the conductivity and the correlation energies.

Rather than computing $\chi^{\lambda}(\mathbf{r},\mathbf{r}',t)$ directly (which in practice requires a huge effort for large systems), a more efficient approach is to obtain $\delta n^{\lambda}(\mathbf{r},t)$ by applying an impulsive perturbation, i.e., $\delta v(\mathbf{r}',t')=\gamma v(\mathbf{r}')\,\delta(t')$ (γ is a small constant with units of time):

$$\delta n^{\lambda}(\mathbf{r},t)=\gamma\int d\mathbf{r}'\chi^{\lambda}(\mathbf{r},\mathbf{r}',t)\,v(\mathbf{r}')\,. \quad (3)$$

Here, $\delta n^{\lambda}(\mathbf{r},t)$ can be computed by applying a perturbation $e^{-i\gamma\hat{v}/\hbar}$ and propagating the perturbed ground state:

$$\delta n_{\gamma}^{\lambda}(\mathbf{r},t)=\left\langle 0_{\lambda}\left|e^{i\gamma\hat{v}/\hbar}\hat{n}^{\lambda}(\mathbf{r},t)\,e^{-i\gamma\hat{v}/\hbar}\right|0_{\lambda}\right\rangle-n_0(\mathbf{r})\,, \quad (4)$$

where $\hat{v}=\int\hat{n}(\mathbf{r}')\,v(\mathbf{r}')\,d\mathbf{r}'$. To see this, expand the right hand side of Eq. (4) to first order in γ: $\delta n_{\gamma}^{\lambda}(\mathbf{r},t)=i\hbar^{-1}\gamma\left\langle 0_{\lambda}\left|\left[\hat{v},\hat{n}^{\lambda}(\mathbf{r},t)\right]\right|0_{\lambda}\right\rangle$ which, when combined with (2) gives Eq. (1).

To obtain the density response $\delta n_{\gamma}^{\lambda}(\mathbf{r},t)$ one needs to solve the many-electron time-dependent Schrödinger equation, which is prohibitive in general. A practical alternative is to use TDDFT. Starting from the KS system of non-interacting electrons having the ground-state density $n_0(\mathbf{r})=2\sum_{j\in occ}|\phi_j(\mathbf{r})|^2$, one perturbs the KS eigenstates $\phi_j(\mathbf{r})$ at $t=0$:

$$\varphi_j(\mathbf{r},t=0)=e^{-i\gamma v(\mathbf{r})/\hbar}\phi_j(\mathbf{r})\,, \quad (5)$$

and then propagates in time according to the TDKS equations

$$i\hbar\frac{\partial\varphi_j(\mathbf{r},t)}{\partial t}=\hat{h}^{\lambda}(t)\,\varphi_j(\mathbf{r},t)\,, \quad (6)$$

where the TDKS Hamiltonian $\hat{h}^{\lambda}(t)$ depends on the screening parameter λ and the propagated density, $n_{\gamma}^{\lambda}(\mathbf{r},t)=2\sum_{j\in occ}|\varphi_j(\mathbf{r},t)|^2$. The density response of Eq. (4) is then obtained from:

$$\int d\mathbf{r}'\chi^{\lambda}(\mathbf{r},\mathbf{r}',t)\,v(\mathbf{r}')=\frac{1}{\gamma}\left(n_{\gamma}^{\lambda}(\mathbf{r},t)-n_{\gamma=0}^{\lambda}(\mathbf{r},t)\right)\equiv\Delta n^{\lambda}(\mathbf{r},t)\,. \quad (7)$$

Eq. (7) simply states that the integral of the susceptibility and a potential $v(\mathbf{r})$ can be computed from the difference between the perturbed and unperturbed densities. This relation holds also for the half Fourier transform quantities ($\tilde{f}(\omega)=\int_0^{\infty}dt\,e^{i\omega t}f(t)$):

$$\int d\mathbf{r}'\tilde{\chi}^{\lambda}(\mathbf{r},\mathbf{r}',\omega)\,v(\mathbf{r}')=\frac{1}{\gamma}\left(\tilde{n}_{\gamma}^{\lambda}(\mathbf{r},\omega)-\tilde{n}_{\gamma=0}^{\lambda}(\mathbf{r},\omega)\right)\equiv\Delta\tilde{n}^{\lambda}(\mathbf{r},\omega)\,. \quad (8)$$

B. Time-Dependent *Stochastic* Density Functional Theory

The stochastic formulation of the density response is identical to the deterministic version outlined above but instead of representing the time-dependent density $n_{\gamma}^{\lambda}(\mathbf{r},t)$ as a sum over all occupied orbital densities ($|\varphi_j(\mathbf{r},t)|^2$) we represent it as an average over the densities of stochastic orbitals $\xi_j(\mathbf{r},t)$.[27] Each stochastic orbital is first projected onto the occupied space and then propagated in time. The advantage of the proposed approach is immediately clear: If the number of stochastic orbitals needed to converge the results does not increase with the system size N, the scaling of the approach is linear with N (rather than quadratic for the deterministic version). Perhaps, in certain cases, due to self-averaging,[27] the scaling will even be better than linear, since the number of stochastic orbitals required to converge the results to a predefined tolerance may decrease with the system size.

The stochastic TDDFT (TDsDFT) procedure is outlined as follows (for simplicity we use a real-space grid representation, but the approach can be generalized to plane-waves or other basis sets):

1. Generate N_{ζ} stochastic orbitals $\zeta_j(\mathbf{r})=e^{i\theta_j(\mathbf{r})}/\sqrt{\delta V}$, where $\theta_j(\mathbf{r})$ is a uniform random variable in the range $[0,2\pi]$, δV is the volume element of the grid, and $j=1,\ldots,N_{\zeta}$. Here, N_{ζ} is typically much smaller than the number of total occupied orbitals (more details below). The stochastic orbitals obey the relation $\mathbf{1}=\langle|\zeta\rangle\langle\zeta|\rangle_{\zeta}$ where $\langle\cdots\rangle_{\zeta}$ denotes a statistical average over ζ.

2. Project each stochastic orbital $\zeta_j(\mathbf{r})$ onto the occupied space: $|\xi_j\rangle\equiv\sqrt{\hat{\theta}_{\beta}}\,|\zeta_j\rangle$, where $\theta_{\beta}(x)=\frac{1}{2}\mathrm{erfc}(\beta(\mu-x))$ is a smooth representation of the Heaviside step function[27] and μ is the chemical potential. The action of $\sqrt{\hat{\theta}_{\beta}}$ is performed using a

suitable expansion in terms of Chebyshev polynomials[32] in the static Hamiltonian with coefficients that depend on μ and β.

3. As in the deterministic case, apply a perturbation at $t=0$: $\xi_j(\mathbf{r},t=0)=e^{-i\gamma v(\mathbf{r})/\hbar}\xi_j(\mathbf{r})$ and propagate the orbitals according to the adiabatic stochastic TDKS equations:

$$i\hbar\frac{\partial\xi_j(\mathbf{r},t)}{\partial t}=\hat{h}^{\lambda}(t)\,\xi_j(\mathbf{r},t), \tag{9}$$

with $\hat{h}^{\lambda}(t) = \hat{h}_{KS} + v^{\lambda}_{HXC}\left[n^{\lambda}_{\gamma}(t)\right](\mathbf{r}) - v^{\lambda}_{HXC}\left[n^{\lambda}_{\gamma}(0)\right](\mathbf{r})$ and

$$v^{\lambda}_{HXC}[n](\mathbf{r})=\lambda\int d\mathbf{r}'\frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}+v^{\lambda}_{XC}(n(\mathbf{r})), \tag{10}$$

where $v^{\lambda}_{XC}(n(\mathbf{r}))$ is the local density (or semi-local) approximation for the exchange correlation potential. For convergence reasons $\hat{h}_{KS}$ is obtained with a rather large number of stochastic orbitals using the sDFT[27] (or its more efficient version, embedded fragment sDFT)[28] and is fixed for the entire propagation. The difference term $v^{\lambda}_{HXC}\left[n^{\lambda}_{\gamma}(t)\right](\mathbf{r}) - v^{\lambda}_{HXC}\left[n^{\lambda}_{\gamma}(0)\right](\mathbf{r})$ is generated with a relatively small number of stochastic orbitals N_ζ and the density

$$n^{\lambda}_{\gamma}(\mathbf{r},t)=2\left\langle|\xi(\mathbf{r},t)|^2\right\rangle_{\zeta}\approx\frac{2}{N_\zeta}\sum_{j=1}^{N_\zeta}|\xi_j(\mathbf{r},t)|^2 \tag{11}$$

is obtained as an *average* over the stochastic orbital densities.

4. Generate $\Delta n^{\lambda}(\mathbf{r},t) = \frac{1}{\gamma}\left(n^{\lambda}_{\gamma}(\mathbf{r},t) - n^{\lambda}_{\gamma=0}(\mathbf{r},t)\right)$, where γ is a small parameter, typically $10^{-3}-10^{-5}\hbar E_h^{-1}$. We note in passing that for $n^{\lambda}_{\gamma=0}(\mathbf{r},t)$ one has to carry out the full propagation since the unperturbed projected stochastic orbitals ($|\xi_j\rangle$) are not eigenstates of the ground-state Hamiltonian. This propagation is not necessary for the deterministic case.

III. RESULTS

A. TDsDFT Calculation of the Absorption Cross Section

The absorption cross section ($\omega \geq 0$) is given by the imaginary part of

$$\sigma(\omega)=\frac{e^2}{3\epsilon_0 c}\omega\int d\mathbf{r}d\mathbf{r}'\,\mathbf{r}\cdot\tilde{\chi}(\mathbf{r},\mathbf{r}',\omega)\cdot\mathbf{r}'. \tag{12}$$

where c is the speed of light. For simplicity, we assume that the perturbing potential is in the z-direction

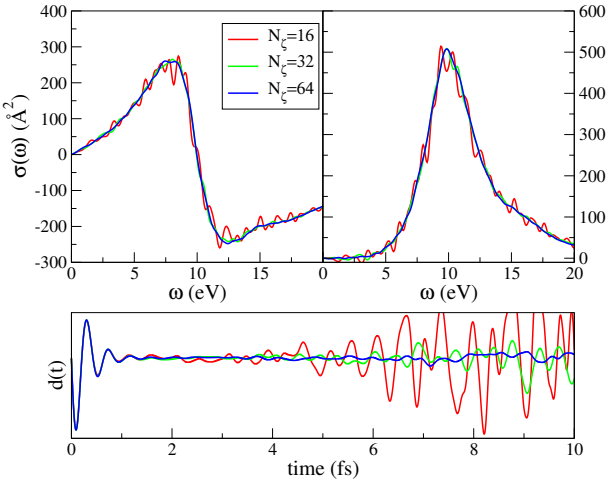

Figure 1. Upper panels: The real (left) and imaginary (right) parts of $\sigma(\omega)$ calculated for $Si_{705}H_{300}$ using $N_\zeta = 16$ (red line), 32 (green line), and 64 (blue line). Lower panel shows the corresponding dipole correlation $d(t)$ as a function of time.

($v(\mathbf{r})=z$) and obtain $\sigma(\omega)$ in Eq. (12) from the Fourier transform of the dipole-dipole correlation function:

$$d_{zz}(t)=\int z\Delta n^{\lambda=1}_{z}(\mathbf{r},t)\,d^3r, \tag{13}$$

where $\Delta n^{\lambda=1}_{z}(\mathbf{r},t)$ is obtained from Eq. (7) and $\sigma(\omega)=\frac{e^2}{\epsilon_0 c}\omega\int_0^\infty dt\,e^{i\omega t}d_{zz}(t)$.

The real and imaginary parts of $\sigma(\omega)$ for $Si_{705}H_{300}$ are plotted in the upper panels of Fig. 1. These and all other results shown in this subsection were generated using the algorithm above within the TDALDA approximation and a grid representation with grid spacing of $\delta x = 0.6a_0$ employing norm-conserving pseudopotentials[33] and image screening methods.[34] We used $\beta = 0.01E_h^{-1}$ to represent the smoothed step-function $\hat{\theta}_\beta$, and a Chebyshev expansion length of 3770 terms. The time-dependent dipole correlation was calculated using a time step of $\delta t = 0.0012$fs up to $t_{max} = 7.5$fs. This signal was multiplied by a Gaussian window function of width 2.5fs and then Fourier transformed to give the absorption cross section.

The right upper panel of Fig. 1 shows the absorption cross-section with a characteristic plasmon frequency of $\sim$ 10eV.[21,35–37] This feature is already captured with $N_\zeta = 16$ stochastic orbitals compared to 1560 occupied orbitals required in the full deterministic TDDFT. It is seen that further increase of N_ζ reduces the statistical fluctuations and provides a handle on the accuracy of the calculation. The convergence of the real part of $\sigma(\omega)$ shown in the upper left panel is similar to its imaginary counterpart.

The calculated dipole correlation $d_{zz}(t)$ is shown in the lower panel of Fig. 1. For these large but finite systems we expect $d_{zz}(t)$ to oscillate and decay to zero at intermediate times followed by recurrences that appear at very

long times (much longer than the timescales shown here). Indeed, the stochastic approximation to $d_{zz}(t)$ oscillates and decays to zero up to a time τ_C, but this is followed by a gradual increase which eventually leads to divergence. This is caused by an instability of the non-linear TDsDFT equations due to the stochastic representation of the time-dependent density. As N_ζ increases and the statistical fluctuations in the density decrease, the divergence onset time τ_C is increased.

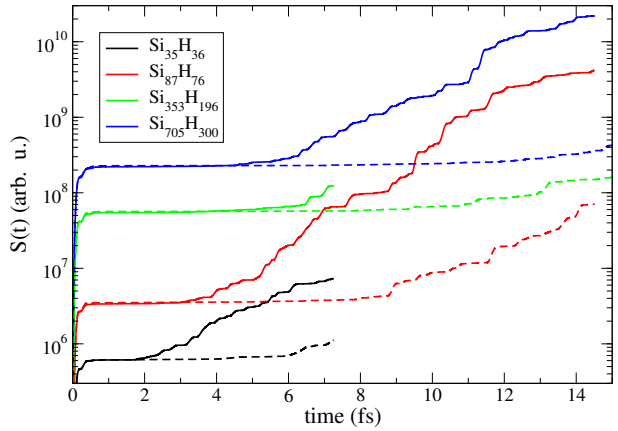

Figure 2. Divergence of the stochastic TDDFT calculation. Shown, the integrated dipole signal $S(t) = \int_0^t d_{zz}(t')^2\, dt'$, where $d_{zz}(t)$ is the calculated dipole correlation as a function of time t, for $N_\zeta = 16$ (solid line) and $N_\zeta = 64$ (dashed line) stochastic orbitals for the different Si NCs. Because of the decay of the dipole correlation the signals reach a plateau after which they diverge sharply due to a nonlinear instability.

In Fig. 2 we plot the integrated dipole signal $S(t) = \int_0^t d_{zz}(t')^2\, dt'$ on a semi-log scale. $S(t)$ provides a clearer measure of τ_C, which is determined as the onset of exponential divergence from the plateau (in practice we take the value of τ_C to be at the middle of the plateau). Two important observations on the onset of the divergence can be noted:

1. τ_C increases for a fixed N_ζ as the system size grows. For $N_\zeta = 16$, τ_C increases from ≈ 1.1fs for $Si_{35}H_{36}$ to ≈ 2.3fs for $Si_{705}H_{300}$. This is a rather moderate, but notable effect, that is a consequence of the so called "self-averaging".[27]

2. τ_C increases with N_ζ for a fixed system size. We find that τ_C roughly scales as $N_\zeta^{1/2}$, namely, an increase of τ_C by 2 requires an increase of N_ζ by 4.

These findings indicate that the number of stochastic orbitals not only determines the level of statistical noise (which scales as $1/\sqrt{N_\zeta}$) but also determines the spectral resolution, given by τ_C^{-1}. To achieve converged results for a fixed cutoff time of $\tau_C = 10$fs, we find that N_ζ decreases from ≈ 1300 for $Si_{35}H_{36}$ to ≈ 230 for $Si_{705}H_{300}$.

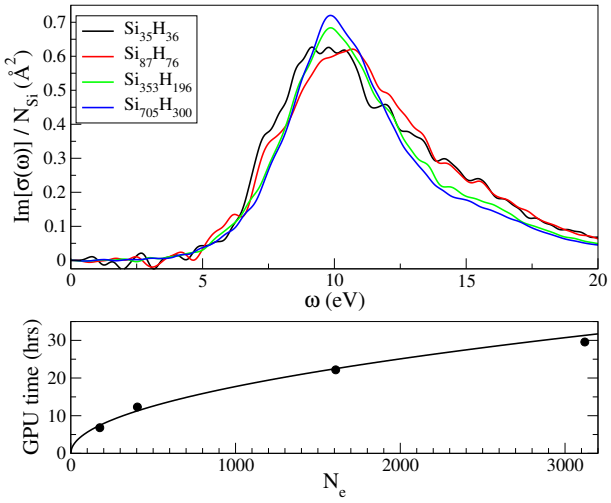

Figure 3. Upper panel: The absorption cross section $\Im\sigma(\omega)$ scaled for size as calculated for several silicon nanocrystals with $N_\zeta = 64$. Lower panel: Extrapolated GPU computational time (T_{GPU}) scaled for a cutoff time of $\tau_C = 10$fs. The line is a power law $T_{GPU} \propto N_e^{0.5}$.

In the upper panel of Fig. 3 we show the absorption cross section for the series of silicon NCs and a fixed number of stochastic orbitals, $N_\zeta = 64$. As the NC size increases the plasmon frequency (peak near 10eV) slightly shifts to lower energies and the width of the plasmon resonance slightly decreases. This is consistent with classical Maxwell equations for which the plasmon frequency depends strongly on the shape but very mildly on the size of the NCs.[38] The statistical fluctuations in the absorption cross section decrease with the system size for a fixed N_ζ, as clearly evident in the figure (most notably at the lower energy range).

The lower panel of Fig. 3 shows the GPU computational time of the approach for a predefined spectral resolution (namely, for converged results up to a fixed cutoff time $\tau_C = 10$fs). Each GPU performs roughly as 3 Intel 3.5GHZ *i*7 third generation quad-core CPUs. Since the number of stochastic orbitals required to converge the results for a fixed time decreases with the system size, the overall scaling of the TDsDFT is better than $O(N_e)$ for the range of sizes studied here, significantly improving the $O(N_e^2)$ scaling of the full deterministic TDDFT. The overall computational effort does depend on the spectral resolution and thus, for small systems or for very high resolution the computational effort of the stochastic approach may exceed that of the full deterministic calculation with all occupied states. But this is certainly not the case for the larger set of NCs studied here, where the wide plasmon resonance dominates the absorption cross section, and thus the spectral features are converged for $\tau_C < 7.5$fs.

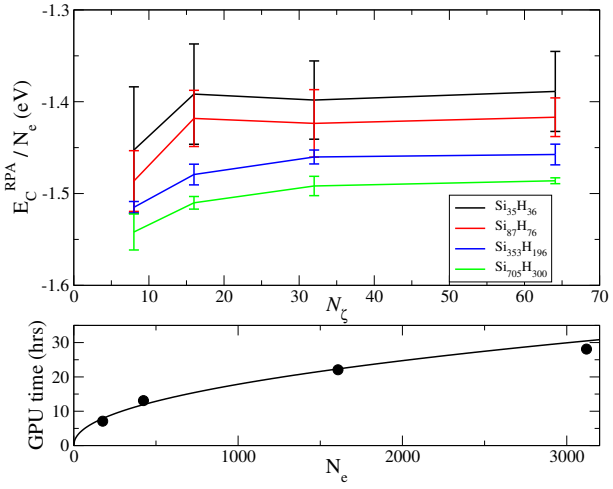

Figure 4. Upper panel: The RPA correlation energy per electron for the silicon NCs as a function of the number of stochastic orbitals N_ζ used to represent the time-dependent density. The error bars are the standard deviation evaluated from 6 statistically independent runs of the algorithm. Lower panel: GPU computational time (T_{GPU}) scaled for a statistical error of 10meV in the total energy per electron. The line is a power law $T_{GPU} \propto N_e^{0.47}$.

B. Stochastic Approach to the Random Phase Approximation Correlation Energy in DFT

The second application of the stochastic TDDFT is for the RPA correlation energy, which is related to $\tilde{\chi}^\lambda(\mathbf{r},\mathbf{r}',\omega)$ by the adiabatic-connection formula:[39,40]

$$E_C^{RPA} = -\frac{\hbar}{2\pi}\Im\int_0^1 d\lambda\int_0^\infty d\omega\int d\mathbf{r}d\mathbf{r}'\times \left(\tilde{\chi}^\lambda(\mathbf{r},\mathbf{r}',\omega)-\tilde{\chi}^0(\mathbf{r},\mathbf{r}',\omega)\right)v_C\left(|\mathbf{r}-\mathbf{r}'|\right), \quad (14)$$

where the integral over λ adiabatically connects the non-interacting density response $\tilde{\chi}^0(\mathbf{r},\mathbf{r}',\omega)$ to the interacting one $\tilde{\chi}^\lambda(\mathbf{r},\mathbf{r}',\omega)$. To proceed, we rewrite Eq. (14) as an average over an additional set of stochastic orbitals $\eta(\mathbf{r}) = e^{i\theta(\mathbf{r})}/\sqrt{\delta V}$

$$E_C^{RPA} = -\frac{\hbar}{2\pi}\Im\int_0^1 d\lambda\int_0^\infty d\omega\int d\mathbf{r}d\mathbf{r}'d\mathbf{r}''\times \left\langle\eta^*(\mathbf{r})\left(\tilde{\chi}^\lambda(\mathbf{r},\mathbf{r}',\omega)-\tilde{\chi}^0(\mathbf{r},\mathbf{r}',\omega)\right)v_C(\mathbf{r}'',\mathbf{r}')\eta(\mathbf{r}'')\right\rangle_\eta. \quad (15)$$

This is done in order to rewrite the perturbation potential as a single-variable potential: $v(\mathbf{r}') = \int d\mathbf{r}' v_C(\mathbf{r}',\mathbf{r})\eta(\mathbf{r}')$, which perturbs the stochastic orbitals at $t=0$: $\xi_j(\mathbf{r},t=0) = e^{-i\gamma v(\mathbf{r})/\hbar}\xi_j(\mathbf{r})$ and from which the density $n_\gamma^\lambda(\mathbf{r},t)$ is computed using Eq. (11). For the propagation of $\xi_j(\mathbf{r},t)$ according to Eq. (9) we set $v_{XC}^\lambda(n(\mathbf{r}))$ to zero, i.e. use the time-dependent Hartree approximation. Using this density and step 4 of the procedure outlined above we compute the density response $\Delta n^\lambda(\mathbf{r},t)$ from which our the RPA correlation energy is calculated:

$$E_C^{RPA} = -\frac{\hbar}{2\pi}\Im\int_0^1 d\lambda\int_0^\infty d\omega\int d\mathbf{r}\times \left\langle\eta^*(\mathbf{r})\left(\Delta\tilde{n}^\lambda(\mathbf{r},\omega)-\Delta\tilde{n}^{\lambda=0}(\mathbf{r},\omega)\right)\right\rangle_\eta. \quad (16)$$

The stochastic formulation for Eq. (16) follows the algorithm described above in Sec. II.

We apply the stochastic RPA formulation to the various silicon NCs studied above. The integration over λ in Eq. (16) was carried out using Gaussian quadrature with 20 sampling points. For each value of λ we used a different set of ζ (for the TDsDFT) and η (for the application of $v(\mathbf{r})$) stochastic orbitals. The TDsDFT total propagation time was 1.5fs with a time step $\delta t = 0.0012$fs, sufficient to converge the RPA correlation energy.

In the upper panel of Fig. 4 we show the calculated the RPA correlation energy per particle for the various silicon NCs as a function of increasing N_ζ, showing convergence as N_ζ increases. The correlation energy per electron grows (in absolute value) with system size, in accordance with our findings in previous studies[41,42] based on a semi-empirical Hamiltonian[43]. The standard deviation (indicated by error bars) evaluated over 6 different runs generally decreases as N_ζ grows for a given system size and also decreases as system size grows for a given value of N_ζ. The magnitude of the error, however, is rather noise due to the small number of independent runs used to estimate it.

The lower panel of Fig. 4 shows the GPU computational time of the approach for a fixed statistical error (estimated as the standard deviation based on the estimate of 20 independent runs) of 10meV. Our previous stochastic formulation of the RPA correlation energy relied on storing all occupied states (memory wise scaled as $O(N_e^2)$) and the computational effort of the RPA stage scaled as $O(N_e^\alpha)$ with $1<\alpha<2$,[42] better than quadratic scaling due to self-averaging. Comparing the current approach with our previous work[42], we find that the present approach shows significant improvements with respect to the computational time and memory requirements. The computational time scales as $O(N_e^{0.47})$ for the range of NCs studied, better than linear scaling for the total RPA correlation energy per electron.

IV. SUMMARY

We have developed a stochastic approach to TDDFT for computing the absorption cross section (via the time-dependent dipole correlation function) and the RPA correlation energy. The core idea of the approach involves time propagation of a set of N_ζ stochastic projected orbitals $\xi_j(\mathbf{r},t)$ according to the time-dependent Kohn-Sham equations. The evolving electron density is exactly represented when $N_\zeta \to \infty$ but the strength of

the method appears when a small number of orbitals $N_\zeta \ll N_e$, where N_e is the number of electrons, is used. Such a truncation produces a statistical fluctuation due to finite sampling. The magnitude of this error is proportional to $1/\sqrt{N_\zeta}$.

The finite sampling error coupled with a nonlinear instability of the time-dependent Kohn-Sham equations produces a catastrophic exponential divergence that becomes noticeable only after a certain propagation time τ_C, which determines the spectral resolution of the approach. The onset of divergence can be controlled by increasing N_ζ and empirically we determined that $\tau_C \propto \sqrt{N_\zeta}$, consistent with the statistical nature of the error.

The TDsDFT was applied to study the absorption cross section and RPA correlation energy for a series of silicon NCs with sizes as large as $N_e \approx 3000$. For this range of NC sizes, the computational time scales sub-linearly, roughly as $O\left(N_e^{1/2}\right)$ for both the absorption cross section and for the RPA correlation energy per electron. For the former, the scaling holds for a given spectral resolution τ_C. Since the computational time is also proportional to $N_\zeta N_e$, one can work backwards to show that $\tau_C \propto \sqrt{N_\zeta N_e^{1/2}}$. For the RPA application, the scaling holds for a given statistical error in the RPA correlation energy per electron.

The developed stochastic TDDFT approach adds another dimension to the arsenal of stochastic electronic structure methods, such as the sDFT[27] (and its more accurate fragmented version)[28] and the sGW.[44] Future work will extend the approach to include exact and screened exchange potentials in order to account for charge-transfer excited states and multiple excitations.

ACKNOWLEDGMENTS

RB and ER are supported by The Israel Science Foundation – FIRST Program (grant No. 1700/14). Y. G. and D. N. are part of the Molecularly Engineered Energy Materials (MEEM), an Energy Frontier Research Center funded by the DOE, Office of Science, Office of Basic Energy Sciences under Award No. de-sc0001342. D. N. also acknowledges support by the National Science Foundation (NSF), Grant .CHE-1112500.

[1] E. Runge and E. K. U. Gross, Phys. Rev. Lett. **52**, 997 (1984).
[2] R. van Leeuwen, Inter. J. Moder. Phys. B **15**, 1969 (2001).
[3] G. Onida, L. Reining, and A. Rubio, Rev. Mod. Phys. **74**, 601 (2002).
[4] M. Marques and E. Gross, Annu. Rev. Phys. Chem. **55**, 427 (2004).
[5] K. Burke, J. Werschnik, and E. K. U. Gross, J. Chem. Phys. **123**, 062206 (2005).
[6] S. Botti, A. Schindlmayr, R. Del Sole, and L. Reining, Rep. Prog. Phys. **70**, 357 (2007).
[7] D. Jacquemin, E. A. Perpete, I. Ciofini, and C. Adamo, Acc. Chem. Res. **42**, 326 (2009).
[8] M. E. Casida, J. Mol. Struct. **914**, 3 (2009).
[9] A. Zangwill and P. Soven, Phys. Rev. A **21**, 1561 (1980).
[10] H. Eshuis, J. E. Bates, and F. Furche, Theor. Chem. Acc. **131**, 1 (2012).
[11] M. Petersilka and E. K. U. Gross, Laser Phys. **9**, 105 (1999).
[12] F. Ceccherini, D. Bauer, and P. Mulser, Laser Part. Beams **18**, 449 (2000).
[13] X. Chu and S. I. Chu, Phys. Rev. A **64**, 063404 (2001).
[14] R. Baer, D. Neuhauser, P. Zdanska, and N. Moiseyev, Phys. Rev. A **68**, 043406 (2003).
[15] K. Nobusada and K. Yabana, Phys. Rev. A **70**, (2004).
[16] A. Castro, H. Appel, M. Oliveira, C. A. Rozzi, X. Andrade, F. Lorenzen, M. A. Marques, E. Gross, and A. Rubio, Phys. Status Solidi B **243**, 2465 (2006).
[17] S. Albrecht, L. Reining, R. Del Sole, and G. Onida, Phys. Rev. Lett. **80**, 4510 (1998).
[18] L. X. Benedict, E. L. Shirley, and R. B. Bohn, Phys. Rev. Lett. **80**, 4514 (1998).
[19] M. Rohlfing and S. G. Louie, Phys. Rev. B **62**, 4927 (2000).
[20] F. Sottile, M. Marsili, V. Olevano, and L. Reining, Phys. Rev. B **76**, 161103 (2007).
[21] L. Ramos, J. Paier, G. Kresse, and F. Bechstedt, Physical Review B **78**, 195423 (2008).
[22] D. Rocca, Y. Ping, R. Gebauer, and G. Galli, Phys. Rev. B **85**, 045116 (2012).
[23] M. E. Casida, "Time-dependent density functional response theory of molecular systems: theory, computational methods, and functionals," in *Recent Developments and Applications in Density Functional Theory*, edited by J. M. Seminario (Elsevier, Amsterdam, 1996) pp. 391–439.
[24] K. Yabana and G. F. Bertsch, Phys. Rev. B **54**, 4484 (1996).
[25] G. F. Bertsch, J. I. Iwata, A. Rubio, and K. Yabana, Phys. Rev. B **62**, 7998 (2000).
[26] R. Baer and D. Neuhauser, J. Chem. Phys. **121**, 9803 (2004).
[27] R. Baer, D. Neuhauser, and E. Rabani, Phys. Rev. Lett. **111**, 106402 (2013).
[28] D. Neuhauser, R. Baer, and E. Rabani, J. Chem. Phys. **141**, 041102 (2014).
[29] G. Giuliani and G. Vignale, *Quantum theory of the electron liquid* (Cambridge University Press, Cambridge, UK ; New York, 2005) pp. xix, 777 p.
[30] A. L. Fetter and J. D. Walecka, *Quantum Thoery of Many Particle Systems* (McGraw-Hill, New York, 1971) p. 299.
[31] R. Kubo, M. Toda, and N. Hashitsume, *Statistical Physics II: Nonequilibrium Statistical Mechanics*, 2nd ed., Springer Series in Solid State Sciences, Vol. 31 (Springer-Verlag, Berlin Heidelberg, 1995).
[32] R. Kosloff, J. Phys. Chem. **92**, 2087 (1988).
[33] N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991).
[34] G. J. Martyna and M. E. Tuckerman, J. Chem. Phys. **110**, 2810 (1999).
[35] J. R. Chelikowsky, L. Kronik, and I. Vasiliev, J. Phys. Condes. Matrer **15**, R1517 (2003).

[36] A. Tsolakidis and R. M. Martin, Physical Review B **71**, 125319 (2005).
[37] M. L. Tiago and J. R. Chelikowsky, Phys. Rev. B **73**, 205334 (2006).
[38] S. Link and M. A. El-Sayed, J. Phys. Chem. B **103**, 4212 (1999).
[39] D. C. Langreth and J. P. Perdew, Solid State Commun. **17**, 1425 (1975).
[40] O. Gunnarsson and B. I. Lundqvist, Phys. Rev. B **13**, 4274 (1976).
[41] D. Neuhauser, E. Rabani, and R. Baer, J. Chem. Theory Comput. **9**, 24 (2013).
[42] D. Neuhauser, E. Rabani, and R. Baer, J. Phys. Chem. Lett. **4**, 1172 (2013).
[43] L. W. Wang and A. Zunger, J. Phys. Chem. **98**, 2158 (1994).
[44] D. Neuhauser, Y. Gao, C. Arntsen, C. Karshenas, E. Rabani, and R. Baer, Phys. Rev. Lett. **113**, 076402 (2014).