Sparse Stochastic Fragmented Exchange for Large-Scale TDDFT Calculations

Mykola Sereda*, Tucker Allen*, and Nadine C Bradbury

Department of Chemistry and Biochemistry, University of California, Los Angeles, California, 90095, USA

Khaled Z Ibrahim

Computer Science Department, Lawrence Berkeley National Laboratory, One Cyclotron road, Berkeley, CA 94720, USA

Daniel Neuhauser*

Department of Chemistry and Biochemistry, and California Nanoscience Institute, UCLA, Los Angeles CA 90095-1569 USA

(Dated: February 25, 2024)

Abstract

In this article, we extend our recently developed sparse stochastic fragmented exchange formalism from ground state DFT (ngH-DFT) to the excited state for linear-response time-dependent Generalized Kohn-Sham DFT (LR-GKS-TDDFT) absorption spectra for systems consisting of thousands of valence electrons within a grid-based/plane-wave representation. Mixed deterministic/fragmented stochastic compression of the exchange kernel, here using the Baer-Neuhauser-Lifsitz (BNL) functional, provides an efficient method for accurate optical spectra. Real-time propagation as well frequency-resolved Casida equation approaches for spectra are presented together, and the method is applied to isolated molecular dye systems.

I. INTRODUCTION

Time-dependent density-functional theory (TDDFT) [1] is the time-dependent extension of density functional theory (DFT), which has been extensively used for predicting photoelectron spectra [2–6], materials screening, optoelectronic device design, [7–11], and calculating nonlinear optical response .[12, 13] A common approximation for the exchange-correlation kernel is the local density approximation (LDA) which is notorious for underestimating the band gap in periodic systems and the ionization potential in finite systems. [14] In time-dependent problems, the commonly used adiabatic LDA approximation leads to incorrect peak positions in absorption spectra, description of charge transfer processes, and cutoff region in higher-harmonic generation spectra. More sophisticated methods, such as the Bethe Salpeter Equation (BSE), TD-CIS(D), or TD-CC2, give a better description of electron correlation and result in a more accurate description of electron dynamics in both the linear response and strong field regime.[15] However, these methods scale poorly with system size, and significant computational or theoretical developments are still needed to address large systems.

On the other hand, range separated hybrid (RSH) functionals with exact exchange in a time-dependent generalized Kohn-Sham (TDGKS) scheme correct the straight line behavior of the ionization potential and quasiparticle bandgap energy while retaining the simplicity of TDDFT with independent orbital propagation.[16] The efficiency and efficacy of TDDFT has

* dxn@ucla.edu

led to significant work on reducing its computational scaling. The methods used to achieve high efficiency within a given TDDFT implementation are often highly dependent on the basis set. In atomic-orbital (AO) basis-sets, efforts to reduce the cost of exact exchange range from density fitting, Cholesky decomposition, resolution of the identity approaches [17–21], and spatial localization methods.[22, 23]. In plane-wave (PW) representations, projection methods have been developed that circumvent the use of virtual orbitals.[24, 25] Adaptively compressed exchange, a low-rank decomposition method, has significantly reduced the cost of exchange and has been applied to the excited state.[26] Diagonal hybrid TDDFT has also been implemented which balances explicit diagonal exchange elements in the Casida matrix with off-diagonal elements approximated by an LDA kernel.[27, 28]

In our previous work, we developed a real-time, grid- based TDGKS algorithm that reduces the steep computational scaling of calculating the exchange by introducing a stochastic resolution of identity that approximates the time-dependent density matrix by summing over a set of stochastic orbitals when applying the exchange operator.[29] An analogous stochastic approach in the frequency domain has also been developed. [30] Both methods introduces a stochastic error that goes as $N_\chi^{-1/2}$ for N_χ stochastic orbitals, so reducing the error requires an increasing number of stochastic orbitals.

Here, we employ the Baer-Neuhauser-Lifshitz (BNL) long-RSH functional with short-range exchange treated locally (or semi-locally), and at long-range full Fock exchange. This functional helps recovers the correct $-1/r$ asymptotic behavior of the exchange potential, necessary for an accurate description of charge-transfer excitations and electron-hole bound states.[16, 29, 31] Previous TDDFT studies with the BNL functional accurately described the vertical excitation energies of a range of acenes, achieving excellent agreement with experimental and coupled cluster values.[32] Further, an optimally-tuned BNL functional was able to correctly predict charge-transfer excitations in light-harvesting molecules.[33, 34]

The new method presented in this article is a different combination of previous techniques that is a radical improvement over previous fully stochastic methods by reducing the computational complexity of the problem and at the same time reducing the stochastic error. First, we work in a select subspace of Kohn-Sham (KS) occupied and unoccupied orbitals, nearest to the Fermi level. This approximation is satisfactory, for example, to resolve the optical gap in the absorption spectrum, however, the large grid size required to calculate the exchange matrix elements for big systems would still be prohibitive. Our

mixed deterministic/fragmented-stochastic approach greatly reduces the explicit number of wavevectors needed to calculate the long-range exact-exchange. We treat the numerically large low-k components of the exchange kernel deterministically and use a sparse stochastic basis for the many numerically small high-k elements. This enables an auxiliary basis approach for electron-repulsion integrals (ERIs) that does not scale with system size.[35]

When combining these techniques, we can extract time-dependent properties of a large system for the cost of a much smaller system. In the following sections, we outline how the sparse fragmented exchange method can be applied to linear-response calculations both in the real-time and frequency domains. The method is applied to naphthalene, fullerene, a Chlorophyll-a (Chla) dye monomer, and hexamer Chlorophyll dye complex found at the reaction center of Photosystem II (RC-PSII). [36, 37]

II. THEORY

Assuming the molecular system of interest responds linearly to the external perturbation of an electric field in a given direction $\hat{r}$, we evaluate the dipole-dipole correlation function through both a real-time propagation (Sec. A) and an eigenvalue problem approach (Sec. B). We go beyond the Tamm-Dancoff approximation (TDA) for the eigenvalue problem, and this yields nearly identical results for both avenues of calculating optical spectra. For completeness, we overview the mixed deterministic/stochastic fragmentation of the Coulomb kernel in Sec. C, and for further detail refer readers to Ref. [35].

A. Hybrid TDDFT in the Valence-Conduction Subspace

The aim of this method is to efficiently solve the TDSE (time-dependent Schrodinger equation),

$$i\partial_t \left|\psi_i\right\rangle = \hat{h}[n(t), \rho(t)] \left|\psi_i\right\rangle , \tag{1}$$

in the TDGKS framework and predict the optical gap in the absorption spectrum.

The starting point for the TDGKS comes from a near-gap Hybrid DFT (ngH-DFT) ground state calculation where the GKS ground state molecular orbitals are represented as

a rotation matrix of the LDA ground state molecular orbitals [35]

$$|\psi_i(t)\rangle = \sum_p C_{pi}(t)\,|\phi_p\rangle\,. \tag{2}$$

In this basis, the GKS Hamiltonian is represented as

$$H_{pq}(t) = \langle\phi_p|\,\hat{h}(t)\,|\phi_q\rangle\,, \tag{3}$$

which is evaluated as,

$$H_{pq}(t) = \langle\phi_p|\,\hat{h}_0 + \delta\hat{v}^{\gamma}[n(t)] + \hat{X}^{\gamma}_{val}[\rho(t)] + \hat{X}^{\gamma}_{core}\,|\phi_q\rangle\,. \tag{4}$$

The range-separation parameter γ is non-empirically obtained by enforcing the ionization potential theorem, further details are provided in Refs. [16, 35]. The time-independent part of the Hamiltonian, $\hat{h}_0$ is

$$\hat{h}_0 = -\frac{1}{2}\hat{\nabla}^2 + \hat{v}^{NL}_{eN} + \hat{v}^{\gamma}[n(t=0)], \tag{5}$$

which includes the Kohn-Sham potential,

$$\hat{v}^{\gamma}[n(t)] = \hat{v}^{local}_{eN} + \int \frac{n(r',t)}{|r-r'|}dr' + \hat{v}^{SR,\gamma}_{XC}[n(t)]. \tag{6}$$

The term, $\hat{v}^{SR,\gamma}_{XC}[n(t)]$ represents the short-range exchange correlation potential within the DFT framework that is calculated with an attenuated short-range Coulomb potential instead. In this work, we use the LDA functional, [38] however, we have the option to use short-range PBE as well.

The next term in the Hamiltonian, $\delta\hat{v}[n(t)]$ is the difference between the short-range Kohn-Sham potential at $t = 0$ and time t,

$$\delta\hat{v}^{\gamma}[n(t)] = \hat{v}^{\gamma}[n(t)] - \hat{v}^{\gamma}[n(t=0)]. \tag{7}$$

Throughout our calculations the core contribution to the density is always approximated

as time-independent, such that,

$$n(r,t) = n_{core}(r) + n_{val}(r,t). \tag{8}$$

The long-range exchange elements are calculated at each time step as,

$$\langle \phi_p | \hat{X}^{\gamma}_{val} | \phi_q \rangle = -\sum_{st} P_{st}(t) \langle \phi_q \phi_s | u^{\gamma} | \phi_t \phi_p \rangle , \tag{9}$$

which is the neargap basis representation of the long-range Coulomb potential,

$$u^{\gamma}(|r-r'|) = \frac{\mathrm{erf}(\gamma|r-r'|)}{|r-r'|}. \tag{10}$$

In order to subvert the prohibitive cost of calculating the exchange matrix elements in Eq. 9, we split the long-range Coulomb potential into two parts, one which we treat deterministically and one that we can accurately compute with a sparse stochastic basis. We discuss the implementation and computational scaling of this method in Section C.

The sum in the density matrix runs over the $N_v \oplus N_c$ subspace. However, there is a correction term, $\hat{X}^{\gamma}_{core}$, that includes the effect of the core states in the exchange. We approximate the effect of the core states as a time-independent scissors correction,

$$\hat{X}^{\gamma}_{core} = \hat{P} d\epsilon_{\gamma,homo} + \hat{Q} d\epsilon_{\gamma.lumo}, \tag{11}$$

where $\hat{P}$ projects into the N_v subspace, $\hat{Q}$ projects into the N_c subspace, and,

$$d\epsilon_{\gamma,a} = \sum_{st \in core} P_{st} \langle \phi_a \phi_s | u^{\gamma} | \phi_t \phi_a \rangle . \tag{12}$$

By inserting the basis representation of the wavefunction in LDA molecular orbitals into the TDSE, and using the orthogonality of the LDA molecular orbitals, we get an ordinary basis set TDSE,

$$i\partial_t C(t) = H(t) C(t). \tag{13}$$

We excite the initial state with a small perturbation,

$$C(t = 0^+) = e^{-i\Delta D} C(t = 0). \tag{14}$$

The dipole matrix elements are,

$$D_{pq} = \langle \phi_p | \, \hat{x} \, | \phi_q \rangle \, . \tag{15}$$

The time-dependent coefficients are propagated using the midpoint rule for the time evolution operator as such,

$$C(t + dt) \simeq e^{-idtH(t+\frac{dt}{2})} C(t). \tag{16}$$

The dipole moment is extracted from the calculation,

$$\mu(t) = \sum_{qp} P_{qp}(t) D_{pq}. \tag{17}$$

The absorption spectra is then generated from the time-dependent dipole moment [15]

$$S(\omega) = \frac{1}{3} Tr[\sigma(\omega)], \tag{18}$$

where,

$$\sigma_{ii}(\omega) = \frac{4\pi\omega}{c} Im[\alpha_{ii}(\omega)]. \tag{19}$$

and,

$$\alpha_{ij}(\omega) = \frac{\mu_i(\omega)}{E_j(\omega)}. \tag{20}$$

Eq. (20) is the ratio of the Fourier transform of the dipole moment in the i-direction, to the transform of the electric field pulse exciting the system in the j-direction, $E_j(\omega)$, which in our case is simply a constant, Δ.

B. $\sigma(\omega)$ via iterative Chebyshev procedure

The frequencies of the TDSE can also be represented as an eigenvalue equation, now known as the Casida equation.[1] For reference, we refer the reader to a traditional derivation including Fock exchange in Ref. [39]. Our fragmented stochastic/deterministic evaluation of exact exchange readily enables the full, non-TDA, calculation of the dipole-dipole correlation

function in frequency space. The Casida equation is expressed

$$\begin{pmatrix} A & B \\ -B & -A \end{pmatrix} \begin{pmatrix} f^+ \\ f^- \end{pmatrix} = \hbar\omega \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} f^+ \\ f^- \end{pmatrix}, \tag{21}$$

with f^+ and f^- referring to the orbitals of the linear response perturbation. In the space of single-particle singlet excitations, the so-called orbital hessians A and B are of the form

$$\begin{aligned} A(ia;jb) &= (\varepsilon_a - \varepsilon_i)\delta_{ij}\delta_{ab} + 2(ia|jb) - (\phi_a\phi_b|u^\gamma(|r-r'|)|\phi_i\phi_j) \\ B(ia;bj) &= 2(ia|bj) - (\phi_a\phi_j|u^\gamma(|r-r'|)|\phi_i\phi_b), \end{aligned} \tag{22}$$

with two-electron integrals (assuming real-valued orbitals)

$$(ia|jb) = \int \phi_i(r)\phi_a(r)|r-r'|^{-1}\phi_j(r')\phi_b(r')dr'dr, \tag{23}$$

so $(ia|jb)$=$(ia|bj)$. In both matrices, the last term contains u^γ. The matrix elements of the long-range exact exchange are

$$\begin{aligned} (\phi_a\phi_b|u^\gamma(|r-r'|)|\phi_i\phi_j) = \\ \int \phi_a(r)\phi_b(r)u^\gamma(|r-r'|)\phi_i(r')\phi_j(r')dr'dr. \end{aligned} \tag{24}$$

In the equations above, we introduced the valence (occupied) $i, j, ...$ and conduction (virtual) $a, b, ...$ energies and associated MOs from a ngH-DFT calculation. [35]

We track the response of the system to an electric field in the x-direction with spinor composed of two dipole matrix elements

$$|\chi\rangle = \begin{pmatrix} \chi^+_{ia} \\ \chi^-_{ia} \end{pmatrix} = \begin{pmatrix} +\langle\phi_a|x|\phi_i\rangle \\ -\langle\phi_a|x|\phi_i\rangle \end{pmatrix} = \begin{pmatrix} f^+_{ia}(t=0) \\ f^-_{ia}(t=0). \end{pmatrix} \tag{25}$$

Application of the Casida matrix, denoted by $\mathcal{H}$, is then

$$\begin{aligned} (\mathcal{H}f^+)_{ia} &= (\varepsilon_a - \varepsilon_i)f^+_{ia} + \langle\phi_a|\delta v_H|\phi_i\rangle \\ &- \langle\phi_a|y^+_i\rangle - \langle\phi_a|z^-_i\rangle, \end{aligned} \tag{26}$$

and for backward transitions,

$$\begin{aligned}(\mathcal{H}f^-)_{ia} &= -(\varepsilon_a - \varepsilon_i)f^-_{ia} - \langle\phi_a|\delta v_H|\phi_i\rangle \\ &+ \langle\phi_a|y^-_i\rangle + \langle\phi_a|z^+_i\rangle,\end{aligned} \tag{27}$$

where $f_{jb} = \langle\phi_b|x|\phi_j\rangle$, and we defined:

$$y^\pm_i(r) = \sum_{jb} f^\pm_{jb}(r)u^\gamma_{ij}(r) \tag{28}$$

$$z^\pm_i(r) = \sum_{jb} f^\pm_{jb}(r)u^\gamma_{ib}(r), \tag{29}$$

where action of u^γ on orbital-pairs is written short-hand $u^\gamma_{ij}(r) \equiv \int u^\gamma(|r-r'|)\phi_i(r')\phi_j(r')dr'$. The Hartree potential is $\delta v_H = \int \frac{\delta n(r')}{|r-r'|}dr'$, and induced density

$$\delta n(r) = 4\sum_{jb}(f^+_{jb} + f^-_{jb})\phi_j(r)\phi_b(r). \tag{30}$$

The iterative Chebyshev approach to obtain the frequency-resolved dipole-dipole correlation function $\sigma(\omega)$ is straightforward.[40] The frequency-resolved absorption spectrum is obtained via

$$\sigma(\omega) \propto \omega\langle\chi|\delta(\mathcal{H} - \omega)|\chi\rangle, \tag{31}$$

where the delta function is evaluated by first finding the Chebyshev residues:

$$R^\pm_n = \langle\tilde{\chi}|T_n(\tilde{\mathcal{H}})|\chi\rangle, \tag{32}$$

where $T_n(\tilde{\mathcal{H}})$ is the n'th Chebyshev polynomial, $\tilde{\mathcal{H}}$ is the scaled Hamiltonian defined within the range $[-1, +1]$, and

$$\tilde{\chi}^\pm_{ia} = \pm\chi^\pm_{ia}. \tag{33}$$

The optical absorption is then obtained from the correlation function as:

$$\sigma(\omega) = \sum_{n=0}^{N_{\text{Cheby}}} c_n(\omega)R_n. \tag{34}$$

For the Chebyshev coefficients $c_n(\omega)$ we use simple smoothly-decaying weights. [40, 41]

Compared with the TDA, our inclusion here in the off-diagonal matrix B and the negative-frequency component f^- (and therefore the addition of "detransitions" of negative frequency) doubles the spectral range, leading to the doubling in the number of required Chebyshev terms, which is typically 1000 now.

C. Deterministic/Fragmented-Stochastic calculation of long-range exchange

Computing the action of $u^\gamma(|r-r'|)$ on orbital pairs, i.e., the exchange of TDHF (time-dependent Hartree Fock), becomes the dominant cost of our spectra calculation. To limit the cost, we use here a mixed deterministic/fragmented-stochastic approach, as developed in Ref. [35], and briefly overviewed below.

As $u^\gamma(r,r') = u^\gamma(|r-r'|)$, we exploit the convolution form of the integrals $u^\gamma_{ij}(r)$ and $u^\gamma_{ib}(r)$:

$$u^\gamma_{ij}(r) = \mathcal{F}^{-1}\Big\{u^\gamma(k)\langle k|\phi_i\phi_j\rangle\Big\} \tag{35}$$

where $\mathcal{F}^{-1}$ denotes an inverse FFT. To reduce the effort in Eq. (35) the interaction $u^\gamma(k)$ is split between long-wavelength, low-k components that are included deterministically, and the remainder high-k terms represented through a sparse stochastic basis, which is of constant size. To achieve this fragmentation of $u^\gamma(k)$, the following identity is introduced:

$$I = \sum_{k_{\rm low}} |k_{\rm low}\rangle\langle k_{\rm low}| + \sum_{k_{\rm high}} |k_{\rm high}\rangle\langle k_{\rm high}|. \tag{36}$$

In this basis, the Coulomb interaction is

$$\begin{aligned} u^\gamma = &\sum_{k_{\rm low}} |k_{\rm low}\rangle u^\gamma(k_{\rm low})\langle k_{\rm low}| \\ &+ \sum_{k_{\rm high}} \sqrt{u^\gamma(k_{\rm high})}|k_{\rm high}\rangle\langle k_{\rm high}|\sqrt{u^\gamma(k_{\rm high})}. \end{aligned} \tag{37}$$

Note that for simplicity we present for the case where all $u^\gamma(k)$ are positive; in practice some elements can be negative due to the use of the Martyna-Tuckerman procedure [42] to avoid grid-reflection effect, so our actual simulations use the general formulation, as detailed in Ref. [35]. We introduce now a sparse stochastic basis α (with N_α members) for the $k_{\rm high}$

space:

$$\alpha(k_{\text{high}}) = \pm\sqrt{\frac{N_{k_{\text{high}}}}{S}} A_\alpha(k_{\text{high}}) \tag{38}$$

where $N_{k_{\text{high}}}$ is the number of high-k terms, S is the length of each fragment, while A_α randomly projects onto a fragment of the k_{high}-space, i.e., is one within the length-S fragment α, and vanishes outside. In theory, for the real-time segment of the paper, we could resample the random projection in the stochastic basis at each time step to reduce the stochastic error. However, in the linear-response regime, the high-k contribution is sufficiently small such that the computational cost of resampling the stochastic basis at each time-step is not worth the reduction in error it would give.

The high-k contribution of u^γ uses then N_α states $|\zeta\rangle$, with components

$$\langle k_{\text{high}}|\zeta\rangle = \sqrt{u^\gamma(k_{\text{high}})}\ \alpha(k_{\text{high}}). \tag{39}$$

The full auxiliary basis with $N_\xi = N_{k_{\text{low}}} + N_\alpha$ components is then defined as

$$|\xi\rangle = \{\sqrt{u^\gamma(k_{\text{low}})}\ |k_{\text{low}}\rangle\} \oplus \{|\zeta\rangle\}, \tag{40}$$

and the interaction is finally a sum of separable terms

$$u^\gamma = \sum_\xi |\xi\rangle\langle\xi|. \tag{41}$$

III. RESULTS

With the TDGKS method, we benchmark on naphthalene and fullerene and then focus our study on a set of molecular dye systems. A chlorophyll a monomer (Chla) and chlorophyll hexamer reaction-center complex found at the center Photosystem II (RC-PSII) are studied. The Chla model has a methyl acetate ligand in place of the phytyl chain, and we mention the optimized coordinates for both dye systems are taken from Ref. [37]. For the Chla system, we focus on the visible wavelength region Q absorption bands that are mostly represented by the HOMO to LUMO transition. These excitations are characteristic of the magnesium-center metalloporphyin ring, and depend on the direction of laser polarization. [43]

The linear-response calculation requires input from a prior ngH-DFT (neargap Hybrid

TABLE I. HOMO-LUMO gaps for naphthalene, fullerene, chlorophyll a (Chla), and 476 atom Photosystem II hexamer dye reaction center (RC-PSII). Also shown are number of grid-points, the total number of occupied states, the maximum numbers valence and conduction states, and the range-separation parameter for each system. All energies are in eV.

System	N_x	N_y	N_z	N_o	$N_v^{\max}$	$N_c^{\max}$	Optimal γ (Bohr^{-1})	LDA-DFT	ngH-DFT
Naphthalene	48	44	24	24	24	104	0.285	3.34	8.63
Fullerene	60	60	60	120	120	480	0.189	1.63	5.42
Chla	84	76	64	116	116	396	0.160	1.40	4.37
RC-PSII	120	148	128	660	200	400	0.120	1.23	3.82

System	Optical gap	Reference
Naphthalene	5.00	4.52
Fullerene	3.33	3.30
Chla	2.01, 2.20	1.99, 2.30
RC-PSII	2.20	1.95

TABLE II. Optical gaps (first peak in spectrum) of naphthalene, fullerene, Chla, and RC-PSII using $N_v^{\max}$:$N_c^{\max}$ states. Q_y,Q_x peaks are given for Chla. Comparison to available literature values are included for dye systems. All energies in eV. Reference values for naphthalene and fullerene are from LR-TDDFT calculations with NWChem software Ref. [44], Chla and RC-PSII are values from Ref. [36, 37].

DFT) calculation. Table I provides necessary parameters for the ground-state calculations of the four systems. The active-space, represented by N_v:N_c, denotes the number of near Fermi-level valence and conduction bands that are explicitly included in both the hybrid DFT calculation and subsequent TD calculation. The spectra obtained for the naphthalene, Chla, and RC-PSII are through the frequency-resolved Chebyshev approach. The resolution of the peaks correspond to a Gaussian broadening of 0.2 eV at half-width. The resolution is fixed by setting the spectral width to number of Chebyshev terms, $\delta\mathcal{H}/N_{\text{Cheby}}$, constant. For fullerene, we used the real-time formalism to obtain the optical gap. In addition, we also performed the same calculations on Chla and naphthalene to ensure that the two methods match exactly. For all of the real-time simulations we used a time step of $dt = 0.25\ a.u.$ for a total time steps of $nt = 5000$.

In Table II, the optical gaps of the naphthalene and fullerene systems are compared with those calculated with the NWChem software package. All optical gaps using our method correspond to the first peak in the absorption spectrum. The NWChem optical gaps are obtained by selecting the lowest root excitation energy with non-zero dipole oscillator

strength. The LR-TDDFT NWChem calculations use an aug-cc-pvdz Gaussian basis set and BNL exchange-correlation functional. The same range-separation parameters presented in Table I are used in the NWChem calculations. Good agreement is obtained between our new method and the established software for these two hydrocarbon systems.

For naphthalene, fullerene, and RC-PSII we present an averaged spectrum, while for Chla the peaks are specifically chosen from the x- and y- polarized spectra. Figures III and III show convergence of the optical gaps of naphthalene and Chla with respect to the explicit number of neargap states included in the spectral calculation.

The convergence of the naphthalene optical gap, Fig. 1, depends more heavily on the inclusion of the virtual orbitals as opposed to the occupied orbitals. However for Chla, Fig. 2, the convergence of the Q_x and Q_y peaks requires that all occupied states and a large number of virtual states are included. For these two systems, we found the optical gap peaks to be converged with the maximum number of conduction states provided in Table I, and a larger conduction basis did not influence the optical gap peak position. In Fig. III, while all occupied orbitals and a large number of conduction states are necessary for convergence of the Q_x and Q_y peaks, the interband distance between the peaks is well-converged even with the limited active-space of $N_v = 20{:}N_c = 40$. Fig. III shows the absorption spectrum of RC-PSII as well as convergence of the optical gap. The optical gap of this much larger system, in contrast with the smaller systems studied, more readily converges with a fraction of valence-conduction states. Similarly, Fig. III shows that the optical gap of the fullerene molecule (C_{60}) is converged within an error 0.025 eV for the set of results ranging from $N_v = 120{:}N_c = 480$ to $N_v = 60{:}N_c = 120$. The dipole signal for $N_v = 120{:}N_c = 480$ and $N_v = 60{:}N_c = 120$ is presented in Fig. III and we see that the dipole signal for the smaller subspace accurately captures the lower frequency oscillations even though the signals do not match exactly.

With regards to stochastic error, we emphasize that stochastic vectors are only introduced for computing part of the exchange interaction, and all other calculations involved are done deterministically. The use of a sparse stochastic basis for the high-k component of the exchange interaction, in both the preparation of the ground-state GKS orbitals and the linear-response spectra calculation, introduces a very small stochastic error. The exchange kernel takes the form $v_\gamma(k) \propto \exp(-k^2/4\gamma^2)/k^2$, which, for $k_{\text{cut}} = 1.8, 1.1, 0.9, 0.5$ a.u. (in order of increasing system size), and γ values provided in Table I, has very numerically

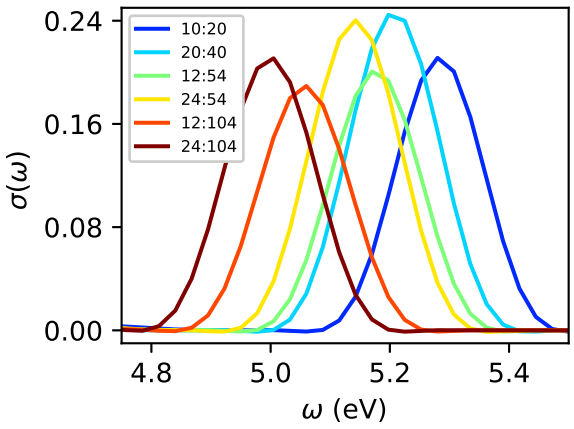

FIG. 1. Convergence of optical gap of naphthalene with respect to N_v:N_c. Casida approach used (absorption cross-section in arbitrary units).

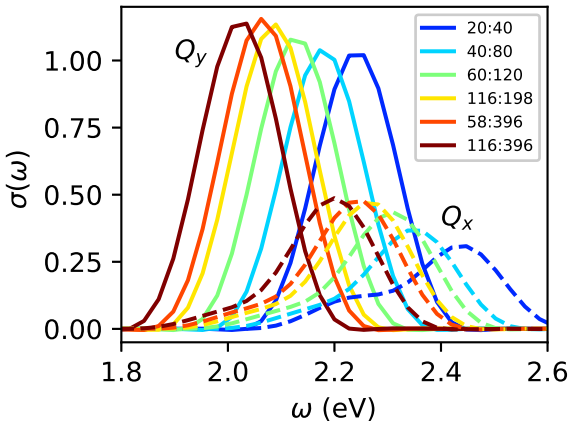

FIG. 2. Convergence of the Q_x and Q_y optical absorption bands (via Casida equation) of Chlorophyll a with respect to active space size N_v:N_c (Subscript x denotes laser polarization in x-direction, y in y-direction, absorption cross-section in arbitrary units).

small elements for the high-k space. The chosen k_{cut} values correspond to $N_{k_{\text{low}}} \simeq 5,000$ low-k elements. Each system uses $N_\alpha = 5,000$ sparse vectors for the high-k space. So an overall auxiliary basis of $N_\xi \simeq 10,000$ is used, irrespective of system-size. We find the stochastic error in our TDDFT calculations is at least an order of magnitude less than the error associated with the choice of N_v and N_c. A more detailed breakdown of the stochastic error is given in Ref. [35]

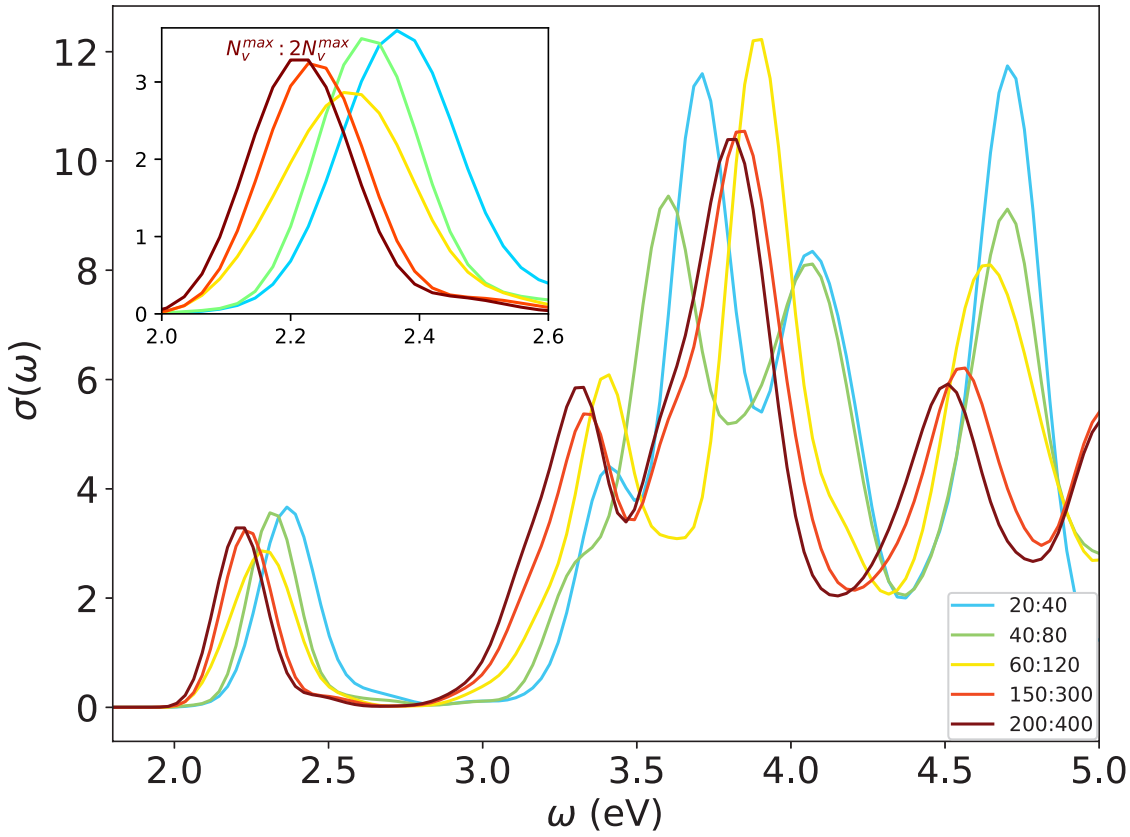

FIG. 3. Absorption spectra of RC-PSII with respect to active space size N_v:N_c. Shown is average $\sigma(\omega)$ (in arbitrary units) of polarization's in x, y, and z directions. The inset shows convergence of optical gap with respect to N_v:N_c. Black curve in both figures detail largest active space used $N_v = 200$:$N_c = 400$. Casida equation approach used for spectra.

IV. DISCUSSION

We have presented a GKS-TDDFT formalism for optical spectra that efficiently calculates long-range exact exchange by a mixed deterministic/fragmented stochastic reciprocal-space grid approach. Both real-time and frequency domain formulations have been investigated.

Traditionally, both real-time and frequency domain solutions to linear-response scale with the number of molecular orbitals. In the frequency domain, the cost to diagonalize the full LR-TDDFT matrix scales as N^6, where N is the total number of molecular orbitals. Various methods exist that extract only a specified number of poles in the spectrum and reduce the scaling to $\sim (N^4 - N^5)$. [45] We use a Chebyshev expansion to iteratively solve for the absorption spectra in frequency space. The real-time solution requires the propagation of the orbitals in either the MO or AO basis. [15] Both of these methods require some method to deal with the overlap matrix. In our real-time code, we have the added benefit of building

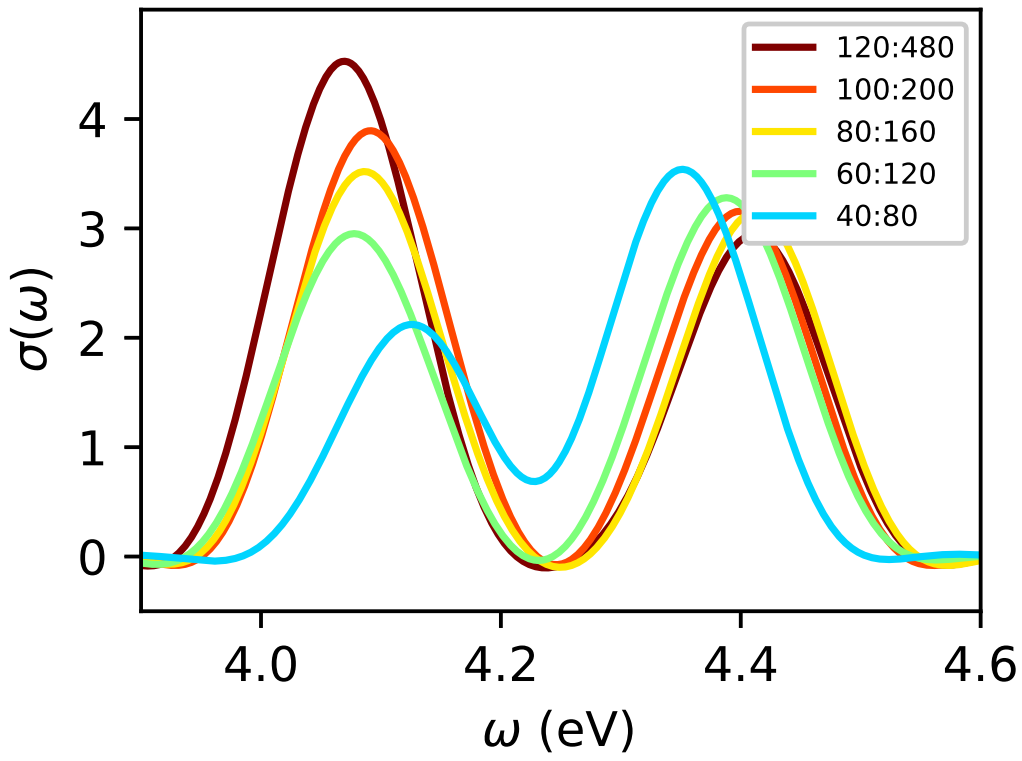

FIG. 4. Absorption spectra of fullerene with respect to active space size N_v:N_c calculated using RT-TDDFT (absorption cross-section in arbitrary units). WILL CHANGE TO REFLECT 3.3 eV optical gap

the Hamiltonian in the same basis that we propagate, since our neargap basis is orthogonal. With our method, we are able to converge the optical gap with a significantly reduced number of occupied MOs. For the RC-PSII, $N_v = 150$ rather than the total $N_o = 660$ can be used, and for fullerene from $N_o = 120$ to $N_v = 60$. The ratio of neargap basis functions relative to the number of molecular orbitals in the fullerene and RC-PSII is significantly smaller when compared to their smaller counterparts, naphthalene and Chla. As expected, the importance of neargap basis functions decreases relative to the number of occupied electrons as system size grows as the chemistry of interest becomes increasingly concentrated near the Fermi energy.

The computational scaling of calculating the exchange matrix elements is reduced from N^4 scaling to $\sim M^2 N_{val} N_\xi$, where $N_\xi = N_{k_{\text{low}}} + N_\alpha$ and M is the full subspace, $M = N_v + N_c$. [35] For our optical gap calculations we use a N_ξ that is approximately constant, and the scaling of N_v and M does not linearly increase with number of total molecular orbitals

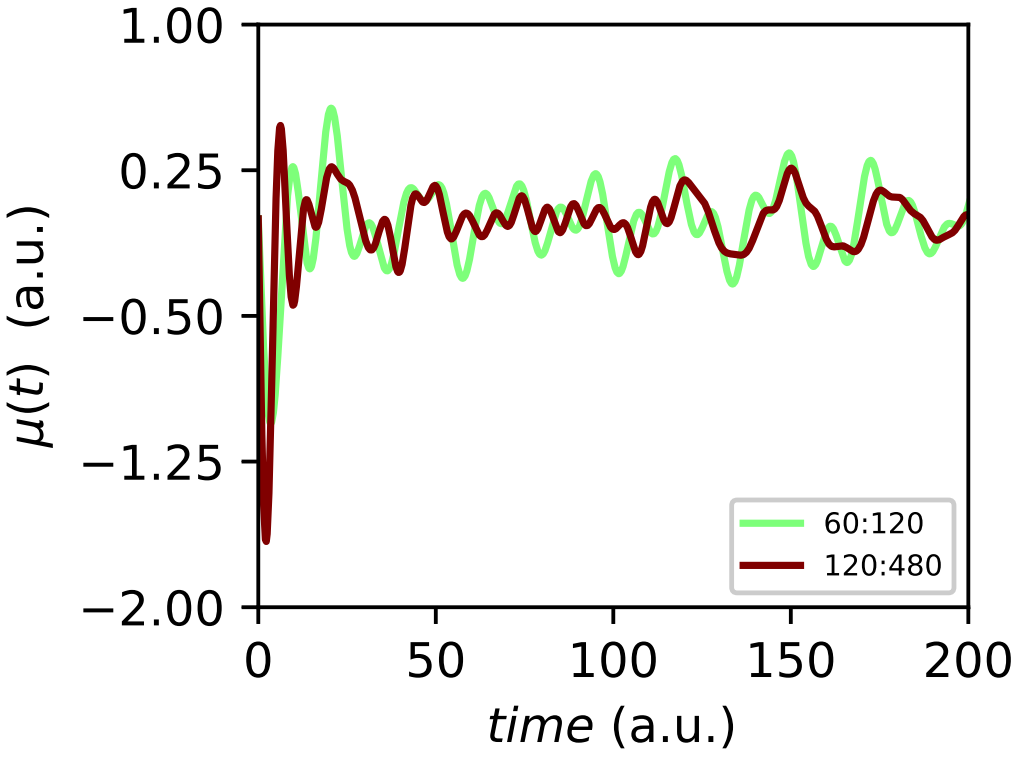

FIG. 5. Dipole signal comparison for fullerene using $N_v = 120$:$N_c = 480$ and $N_v = 60$:$N_c = 120$.

in the system. The combination of all of our techniques culminates in a method that is particularly well suited to extract the optical properties for systems with a high number of occupied electrons.

In our naphthalene and Chla simulations we find that more conduction states are needed, compared to valence states, to converge the optical gap. In the case of missing occupied electrons, the core correction to the exchange accounts for much the effect of the occupied states not included in the valence subspace by shifts the peak positions with a rigid scissor shift. There is no such equivalent formalism to address the effect of the missing high-lying conduction states. In the future, we will attempt to improve the convergence with respect to the conduction states by searching for a rotation on the conduction subspace that orders the orbitals by increasing magnitude of the information contained in the optical gap transition.

The convergence with respect to the selected active-space will be significantly improved when extending this code to the GW-BSE method because the screened Coulomb interaction is much weaker than the exchange interaction in the current TDGKS formalism. The

inclusion of exact exchange in this method allows a better description of charge-transfer excitations, however one needs explicit inclusion of the electron-hole attractive interaction for a more accurate absorption spectrum. The method is easily extendable to our frequency-resolved BSE code with the screened Coulomb interaction replacing the long-RSH in the exchange term. Our recent formulation of the stochastic BSE exactly splits the static screened Coulomb interaction $W(r, r')$ to a fitted exchange interaction and a small difference that is sampled stochastically.[40] This makes the bulk of computational effort calculating convolutions, which we have substantially improved the efficiency of in the present work.

In LR-TDDFT, as well as BSE calculations, frequency-domain calculations often employ the Tamm-Dancoff approximation (TDA).[46] Two key reasons being the unfavorable computational cost of the full non-Hermitian eigenvalue problem as well as the fact that mixing of upward and downward vertical transitions often makes minor changes in the final spectra and excitation energies. However, the TDA in BSE breaks down for carbon nanotubes, π conjugated systems, and other low-dimensional materials. [47, 48] Using our current methods to go beyond the TDA in BSE calculations will give absorption spectra that in addition to electron-hole interaction include resonant-antiresonant coupling.

In addition to the extension of this LR-TDDFT code to the BSE, we detail future developments/extensions of the presented method.

Fragmented stochastic exchange can be formulated within an orthogonal-projector augmented waves (OPAW) framework. This would extend our recent LDA-OPAW-TDDFT to hybrid functionals with long-range exact exchange. [49] OPAW enables use of coarser grids and a lower kinetic-energy cutoff compared to norm-conserving pseudopotentials, which will unlock even larger system-size calculations.

Another pursuit for the fragmented-stochastic GKS-TDDFT approach will be the possibility for excited state nuclear gradients. Previous work computing ionic forces in TDDFT were limited to LDA functionals, and the presented method significantly reduces the cost of exchange making long-RSH functionals more accessible. [50] We also mention recent work done with local hybrids.[51].

Going beyond the linear-response regime, the real-time code is particularly suited for simulating strong field phenomena such as strong field ionization and high harmonic generation. Even though most practical calculations are often done on small systems, the system of interest is still large due to the large grid space required to resolve the electron density

when it is far from equilibrium. We expect to see a different pattern of behavior in converging both the $N_v \oplus N_c$ subspace as well as the parameters in splitting the long-range Coulomb interaction. We are also interested to see the performance of this code to calculate the hyperpolarizability of molecules.

ACKNOWLEDGEMENTS

This work is supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Scientific Discovery through Advanced Computing (SciDAC) program under Award Number DE-SC0022198. Computational resources for simulations were provided by the Expanse cluster at San Diego Supercomputer Center through allocation CHE220086 from the Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS) program. [52] NCB acknowledges the NSF Graduate Research Fellowship Program under grant DGE-2034835.

---

[1] M. E. Casida, Time-Dependent Density Functional Response Theory for Molecules, in *Recent Advances in Computational Chemistry*, Vol. 1 (WORLD SCIENTIFIC, 1995) pp. 155–192.

[2] C. Jamorski, M. E. Casida, and D. R. Salahub, Dynamic polarizabilities and excitation spectra from a molecular implementation of time-dependent density-functional response theory: N2 as a case study, The Journal of Chemical Physics **104**, 5134 (1996).

[3] K. Yabana and G. F. Bertsch, Time-dependent local-density approximation in real time: Application to conjugated molecules, International Journal of Quantum Chemistry **75**, 55 (1999).

[4] K. Yabana and G. Bertsch, Optical response of small carbon clusters, Zeitschrift fr Physik D Atoms, Molecules and Clusters **42**, 219 (1997).

[5] K. Lopata, B. E. Van Kuiken, M. Khalil, and N. Govind, Linear-Response and Real-Time Time-Dependent Density Functional Theory Studies of Core-Level Near-Edge X-Ray Absorption, Journal of Chemical Theory and Computation **8**, 3284 (2012).

[6] S. DeBeer George, T. Petrenko, and F. Neese, Time-dependent density functional calculations of ligand K-edge X-ray absorption spectra, Inorganica Chimica Acta **361**, 965 (2008).

[7] J. Lermé, B. Palpant, B. Prével, E. Cottancin, M. Pellarin, M. Treilleux, J. Vialle, A. Perez, and M. Broyer, Optical properties of gold metal clusters: A time-dependent local-density-approximation investigation, The European Physical Journal D - Atomic, Molecular and Optical Physics **4**, 95 (1998).

[8] X. Qian, J. Li, X. Lin, and S. Yip, Time-dependent density functional theory with ultrasoft pseudopotentials: Real-time electron propagation across a molecular junction, Physical Review B **73**, 035408 (2006).

[9] R. Baer and D. Neuhauser, Ab initio electrical conductance of a molecular wire, International Journal of Quantum Chemistry **91**, 524 (2003).

[10] R. Baer and R. Gould, A method for *ab initio* nonlinear electron-density evolution, The Journal of Chemical Physics **114**, 3385 (2001).

[11] T. Zelovich, L. Kronik, and O. Hod, State Representation Approach for Atomistic Time-Dependent Transport Calculations in Molecular Junctions, Journal of Chemical Theory and Computation **10**, 2927 (2014).

[12] F. Ding, B. E. Van Kuiken, B. E. Eichinger, and X. Li, An efficient method for calculating dynamical hyperpolarizabilities using real-time time-dependent density functional theory, The Journal of Chemical Physics **138**, 064104 (2013).

[13] S. M. Parker, D. Rappoport, and F. Furche, Quadratic Response Properties from TDDFT: Trials and Tribulations, Journal of Chemical Theory and Computation **14**, 807 (2018).

[14] J. P. Perdew, Density functional theory and the band gap problem, International Journal of Quantum Chemistry **28**, 497 (2009).

[15] X. Li, N. Govind, C. Isborn, A. E. DePrince, and K. Lopata, Real-Time Time-Dependent Electronic Structure Theory, Chemical Reviews **120**, 9951 (2020).

[16] R. Baer, E. Livshits, and U. Salzner, Tuned range-separated hybrids in density functional theory, Annual Review of Physical Chemistry **61**, 85 (2010).

[17] F. Aquilante, T. B. Pedersen, and R. Lindh, Low-cost evaluation of the exchange Fock matrix from Cholesky and density fitting representations of the electron repulsion integrals, The Journal of Chemical Physics **126**, 194106 (2007).

[18] S. F. Manzer, E. Epifanovsky, and M. Head-Gordon, Efficient implementation of the pair atomic resolution of the identity approximation for exact exchange for hybrid and range-separated density functionals, Journal of Chemical Theory and Computation **11**, 518 (2015).

[19] X. Qin, J. Liu, W. Hu, and J. Yang, Interpolative separable density fitting decomposition for accelerating Hartree–Fock exchange calculations within numerical atomic orbitals, The Journal of Physical Chemistry A **124**, 5664 (2020).

[20] M. Medves, G. Fronzoni, and M. Stener, Optimization of density fitting auxiliary Slater-type basis functions for time-dependent density functional theory, Journal of Computational Chemistry **43**, 1923 (2022).

[21] S. Sharma, A. F. White, and G. Beylkin, Fast exchange with Gaussian basis set using robust pseudospectral method, Journal of Chemical Theory and Computation **18**, 7306 (2022).

[22] J. Liu and J. Herbert, An efficient and accurate approximation to time-dependent density functional theory for systems of weakly coupled monomers, The Journal of chemical physics **143**, 034106 (2015).

[23] H.-Y. Ko, B. Santra, and R. A. J. DiStasio, Enabling large-scale condensed-phase hybrid density functional theory-based ab initio molecular dynamics ii: Extensions to the isobaric–isoenthalpic and isobaric–isothermal ensembles, Journal of Chemical Theory and Computation **17**, 7789 (2021).

[24] J. Hutter, Excited state nuclear forces from the Tamm-Dancoff approximation to time-dependent density functional theory within the plane wave basis set framework, J. Chem. Phys. **118**, 3928 (2003).

[25] N. M. Boffi, M. Jain, and A. Natan, Efficient computation of the hartree–fock exchange in real-space with projection operators, Journal of Chemical Theory and Computation **12**, 3614 (2016).

[26] J. Liu, W. Hu, and J. Yang, Accelerating linear-response time-dependent hybrid density functional theory with low-rank decomposition techniques in the plane-wave basis, Journal of Chemical Theory and Computation **18**, 6713 (2022).

[27] M. Medves, L. Sementa, D. Toffoli, G. Fronzoni, A. Fortunelli, and M. Stener, An efficient hybrid scheme for time dependent density functional theory, The Journal of Chemical Physics **152**, 184104 (2020).

[28] P. D'Antoni, M. Medves, D. Toffoli, A. Fortunelli, M. Stener, and L. Visscher, A resolution of identity technique to speed up TDDFT with hybrid functionals: Implementation and application to the magic cluster series au8n+4(sc6h5)4n+8 (n = 3–6), The Journal of Physical Chemistry A **127**, 9244 (2023).

[29] V. Vlcek, R. Baer, and D. Neuhauser, Stochastic time-dependent DFT with optimally tuned range-separated hybrids: Application to excitonic effects in large phosphorene sheets, J. Chem. Phys. **150**, 184118 (2019).

[30] X. Zhang, G. Lu, R. Baer, E. Rabani, and D. Neuhauser, Linear-response time-dependent density functional theory with stochastic range-separated hybrids, Journal of Chemical Theory and Computation **16**, 1064 (2020).

[31] R. Baer and D. Neuhauser, Density functional theory with correct long-range asymptotic behavior, Phys. Rev. Lett. **94**, 043002 (2005).

[32] K. Lopata, R. Reslan, M. Kowalska, D. Neuhauser, N. Govind, and K. Kowalski, Excited-state studies of polyacenes: A comparative picture using EOMCCSD, CR-EOMCCSD(T), Range-Separated (LR/RT)-TDDFT, TD-PM3, and TD-ZINDO, Journal of Chemical Theory and Computation **7**, 3686 (2011).

[33] A. Karolewski, T. Stein, R. Baer, and S. Kümmel, Communication: Tailoring the optical gap in light-harvesting molecules, The Journal of Chemical Physics **134**, 151101 (2011).

[34] L. Kronik, T. Stein, S. Refaely-Abramson, and R. Baer, Excitation gaps of finite-sized systems from optimally tuned range-separated hybrid functionals, Journal of Chemical Theory and Computation **8**, 1515 (2012).

[35] N. C. Bradbury, T. Allen, M. Nguyen, and D. Neuhauser, Deterministic/fragmented-stochastic exchange for large-scale hybrid DFT calculations, Journal of Chemical Theory and Computation **19**, 9239 (2023).

[36] E. Gruber, C. Kjær, S. B. Nielsen, and L. H. Andersen, Intrinsic photophysics of light-harvesting charge-tagged chlorophyll a and b pigments, Chemistry – A European Journal **25**, 9153 (2019).

[37] A. Förster and L. Visscher, Quasiparticle self-consistent GW-Bethe-Salpeter equation calculations for large chromophoric systems, Journal of Chemical Theory and Computation **18**, 6779 (2022).

[38] J. P. Perdew and Y. Wang, Accurate and simple analytic representation of the electron-gas correlation energy, Phys. Rev. B **45**, 13244 (1992).

[39] J. W. Negele, The mean-field theory of nuclear structure and dynamics, Reviews of Modern Physics **54**, 913 (1982).

[40] N. C. Bradbury, T. Allen, M. Nguyen, K. Z. Ibrahim, and D. Neuhauser, Optimized attenuated interaction: Enabling stochastic Bethe–Salpeter spectra for large systems, The Journal of Chemical Physics **158**, 154104 (2023).

[41] A. Weiße, G. Wellein, A. Alvermann, and H. Fehske, The kernel polynomial method, Rev. Mod. Phys. **78**, 275 (2006).

[42] G. J. Martyna and M. E. Tuckerman, A reciprocal space based method for treating long range interactions in ab initio and force-field-based calculations in clusters, The Journal of Chemical Physics **110**, 2810 (1999).

[43] B. Kim and J. Bohandy, Spectroscopy of porphyrins., **2**, 153 (1981).

[44] E. Aprà, E. J. Bylaska, W. A. de Jong, N. Govind, K. Kowalski, T. P. Straatsma, M. Valiev, H. J. J. van Dam, Y. Alexeev, J. Anchell, V. Anisimov, F. W. Aquino, R. Atta-Fynn, J. Autschbach, N. P. Bauman, J. C. Becca, D. E. Bernholdt, K. Bhaskaran-Nair, S. Bogatko, P. Borowski, J. Boschen, J. Brabec, A. Bruner, E. Cauët, Y. Chen, G. N. Chuev, C. J. Cramer, J. Daily, M. J. O. Deegan, J. Dunning, T. H., M. Dupuis, K. G. Dyall, G. I. Fann, S. A. Fischer, A. Fonari, H. Früchtl, L. Gagliardi, J. Garza, N. Gawande, S. Ghosh, K. Glaesemann, A. W. Götz, J. Hammond, V. Helms, E. D. Hermes, K. Hirao, S. Hirata, M. Jacquelin, L. Jensen, B. G. Johnson, H. Jónsson, R. A. Kendall, M. Klemm, R. Kobayashi, V. Konkov, S. Krishnamoorthy, M. Krishnan, Z. Lin, R. D. Lins, R. J. Littlefield, A. J. Logsdail, K. Lopata, W. Ma, A. V. Marenich, J. Martin del Campo, D. Mejia-Rodriguez, J. E. Moore, J. M. Mullin, T. Nakajima, D. R. Nascimento, J. A. Nichols, P. J. Nichols, J. Nieplocha, A. Otero-de-la Roza, B. Palmer, A. Panyala, T. Pirojsirikul, B. Peng, R. Peverati, J. Pittner, L. Pollack, R. M. Richard, P. Sadayappan, G. C. Schatz, W. A. Shelton, D. W. Silverstein, D. M. A. Smith, T. A. Soares, D. Song, M. Swart, H. L. Taylor, G. S. Thomas, V. Tipparaju, D. G. Truhlar, K. Tsemekhman, T. Van Voorhis, Á. Vázquez-Mayagoitia, P. Verma, O. Villa, A. Vishnu, K. D. Vogiatzis, D. Wang, J. H. Weare, M. J. Williamson, T. L. Windus, K. Woliński, A. T. Wong, Q. Wu, C. Yang, Q. Yu, M. Zacharias, Z. Zhang, Y. Zhao, and R. J. Harrison, NWChem: Past, present, and future, The Journal of Chemical Physics **152**, 184102 (2020).

[45] J. Brabec, L. Lin, M. Shao, N. Govind, C. Yang, Y. Saad, and E. G. Ng, Efficient Algorithms for Estimating the Absorption Spectrum within Linear Response TDDFT, Journal of Chemical Theory and Computation **11**, 5197 (2015).

[46] S. Hirata and M. Head-Gordon, Time-dependent density functional theory within the Tamm–Dancoff approximation, Chemical Physics Letters **314**, 291 (1999).

[47] Y. Ma, M. Rohlfing, and C. Molteni, Excited states of biological chromophores studied using many-body perturbation theory: Effects of resonant-antiresonant coupling and dynamical screening, Physical Review B **80** (2009).

[48] M. Gruning, A. Marini, and X. Gonze, Exciton-plasmon states in nanoscale materials: Breakdown of the TammDancoff approximation, Nano Letters **9**, 2820 (2009).

[49] M. Nguyen, T. Duong, and D. Neuhauser, Time-dependent density functional theory with the orthogonal projector augmented wave method (2023), arXiv:2312.14179 [physics.chem-ph].

[50] X. Zhang and G. Lu, Subspace formulation of time-dependent density functional theory for large-scale calculations., The Journal of chemical physics **143 6**, 064110 (2015).

[51] R. Grotjahn, F. Furche, and M. Kaupp, Development and implementation of excited-state gradients for local hybrid functionals, Journal of Chemical Theory and Computation **15**, 5508 (2019).

[52] T. J. Boerner, S. Deems, T. R. Furlani, S. L. Knuth, and J. Towns, ACCESS: advancing innovation: NSF's advanced cyberinfrastructure coordination ecosystem: services & support, in *Practice and Experience in Advanced Research Computing* (2023) pp. 173–176.