Sparse-Stochastic Fragmented Exchange for Large-Scale Hybrid Time-Dependent Density Functional Theory Calculations

Mykola Sereda,[§] Tucker Allen,[§] Nadine C. Bradbury, Khaled Z. Ibrahim, and Daniel Neuhauser*

Read Online

ACCESS | 📊 Metrics & More | 📄 Article Recommendations

ABSTRACT: We extend our recently developed sparse-stochastic fragmented exchange formalism for ground-state near-gap hybrid DFT to calculate absorption spectra within linear-response time-dependent generalized Kohn−Sham DFT (LR-GKS-TDDFT) for systems consisting of thousands of valence electrons within a grid-based/plane-wave representation. A mixed deterministic/fragmented-stochastic compression of the exchange kernel, here using long-range explicit exchange functionals, provides an efficient method for accurate optical spectra. Both real-time propagation as well as frequency-resolved Casida-equation-type approaches for spectra are presented, and the method is applied to large molecular dyes.

LR-GKS-TDDFT Spectra for Photosystem II Reaction Center with Sparse-Stochastic Exchange
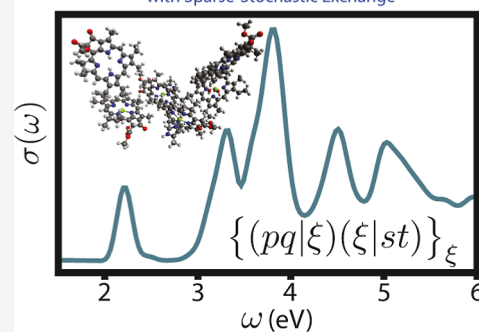
1. INTRODUCTION

Time-dependent density functional theory (TDDFT)[1] has been extensively used for predicting photoelectron spectra,[2−6] material screening, optoelectronic device design,[7−11] and nonlinear optical response.[12,13] A common approximation for the exchange−correlation kernel is the local density approximation (LDA), which in DFT is notorious for underestimating the band gap in periodic compounds and the ionization potential in finite systems.[14] Similarly, in time-dependent calculations, the commonly used adiabatic LDA (ALDA) approximation often gives incorrect absorption peaks and shifted cutoff regions 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 can scale poorly with system size, and significant computational or theoretical developments are still needed to address large systems.

Hybrid functionals in DFT admix exact exchange with an approximate exchange−correlation (XC) functional.[16] Inclusion of exact exchange helps improve band gaps and alleviate the self-repulsion in the Kohn−Sham (KS) potential.[17] Global hybrids, such as PBE0 and B3LYP, include a fractional amount of exchange at all length scales.[18] On the other hand, range-separated hybrid (RSH) functionals partition the Coulomb interaction into a short-range component, treated locally, and a long-range part treated explicitly. The strategy is to correct the asymptotic behavior of the KS potential while maintaining a local (or semilocal) XC potential description at short range.[19,20]

Popular RSHs, such as CAM-B3LYP, use a fixed range−separation parameter, here denoted by γ, to partition the short- and long-range components of the exchange.[21,22] Optimally tuned (OT)-RSHs obtain γ by enforcing a physical condition on the system of interest based on first-principles.[20] OT-RSH functionals recover the correct $-1/r$ asymptotic behavior of the exchange potential, essential for an accurate description of charge-transfer excitations and electron−hole bound states.[17,23] In this work, we use the Baer−Neuhauser−Lifshitz (BNL) OT-RSH functional.[20] For this functional, the range-separation parameter γ is obtained by enforcing piece-wise linearity of the total energy with electron number.[17]
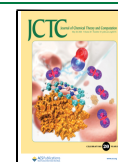
For excited states, ALDA-TDDFT simplifies the time-dependent many-body problem to an independent orbital propagation. Use of a hybrid functional retains this independent orbital propagation and significantly improves the ailments of ALDA. Previous TDGKS studies with the BNL functional accurately described the vertical excitation energies of a range of acenes, achieving excellent agreement with experimental and coupled cluster values.[24] Furthermore, an OT BNL functional was able to correctly predict charge-

transfer excitations in light-harvesting molecules.[25,26] However, this incurs a steep computational scaling.

Methods to reduce the computational scaling within a given TDGKS implementation depend on the choice of basis-set. (We use below the terms TDGKS and hybrid-TDDFT interchangeably.) In atomic-orbital basis sets, efforts to reduce the cost of exact exchange include density fitting, Cholesky decomposition, resolution of the identity approaches,[27−31] and spatial localization methods.[32,33] In plane-wave (PW) and grid-based representations, projection methods have been developed that circumvent the use of virtual orbitals.[34,35] Adaptively compressed exchange, a low-rank decomposition method, has significantly reduced the cost of exchange and has been applied to excited states.[36] 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.[37,38]

In our previous work, we developed a real-time, grid-based TDGKS algorithm that reduces the scaling of exact exchange by a stochastic decomposition of the time-dependent density matrix and the Coulomb kernel.[23] A similar stochastic approach in the frequency domain has also been developed.[39]

Here, we present an alternate approach with very little stochastic error by introducing a mixed deterministic/sparse-stochastic sampling of the Coulomb kernel in k-space and avoiding a stochastic resolution of the density matrix. The numerically large low-k components of the exchange kernel are treated deterministically and the many numerically small high-k elements are represented through a sparse-stochastic basis. The size of this sparse high-k basis becomes constant, irrespective of system size.[40] In addition, we work in a valence-conduction subspace of KS orbitals, nearest to the Fermi level.

When combining these techniques, we extract readily, with modest costs, the time-dependent properties of large systems. In the following sections, we outline how the sparse-fragmented exchange method is applied to linear-response calculations in both the real-time and frequency domains. The method is applied to naphthalene, fullerene, a chlorophyll a (Chla) dye monomer, and a hexamer chlorophyll dye complex found at the reaction center of Photosystem II (RC-PSII).[41,42] The latter system is large (1320 valence electrons) and therefore very expensive for traditional deterministic approaches.

2. THEORY

For a small external perturbation of an electric field in a given direction, taken as $\hat{x}$ here for simplicity, we evaluate the dipole–dipole correlation function through either a real-time propagation (Section 2.1) or a frequency-domain linear-response approach (Section 2.2).

The two approaches each have their own specific benefits. The real-time approach is simple to derive and easily extendable to strong and continuous electric field pulses. On the other hand, an iterative frequency-domain method, as used here, which iterates only a single exciton vector to find directly $\delta(\mathcal{L} - \omega)|\chi\rangle$, is numerically faster and specifically optimized for extracting the low-frequency linear-response absorption spectra. Note that beyond a few hundred atoms, the iterative methods, both frequency-domain and real-time, become more efficient than frequency-domain methods that solve the complete Casida matrix to obtain multiple low-lying excitons of a molecular system.

Since we go beyond the Tamm−Dancoff approximation (TDA) for the frequency-domain approach, the two methods yield essentially identical results for the linear-response optical spectra. For completeness, we overview the mixed deterministic/stochastic fragmentation of the Coulomb kernel in Section 2.3, and for further details, we refer readers to ref 40.

2.1. Hybrid TDDFT in the Valence−Conduction Subspace. The aim of the real-time approach is to efficiently solve the TDKS (time-dependent Kohn−Sham) equation

$$i\partial_t|\psi_i\rangle = \hat{h}|\psi_i\rangle \tag{1}$$

In the GKS framework and predict the optical gap in the absorption spectrum. The algorithm-specific density and density matrix will be defined later in the section.

The starting point comes from a near-gap hybrid DFT (ngH-DFT) ground-state calculation, where the GKS ground-state molecular orbitals (MOs) are represented as a rotation matrix of the LDA ground-state MOs[40]

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

where the basis set is the subset of eigenfunctions of the ground-state LDA KS-DFT Hamiltonian

$$\hat{H}\phi_s = \varepsilon_s \phi_s \tag{3}$$

The MOs, ϕ_p, are divided into four sets of states: N_{core} core, $N_v = N_{occ} - N_{core}$ valence (occupied), N_c conduction (virtual), and the remaining high-lying conduction states. We work in the combined valence and conduction subspace, $N_v \oplus N_c$. The effect of the high-lying conduction subspace is not considered, and the effect of the core states is approximated as a perturbative scissor correction, discussed in eq 15.

We emphasize that our labeling of "core" states, as in ref 40, merely denotes lower-energy MOs that play a lesser role in the description of optical excitations. The division between valence and core states is done arbitrarily, so the number of valence states is a numerical convergence parameter. All states here share the same PW (or equivalently grid) representation, although the approach is also applicable if the underlying local-DFT MO ϕ_p were given in terms of a localized basis set.

Throughout this paper, the indices p, q,... run over the combined $M = N_v + N_c$ near-gap states, i, j,... refer to valence orbitals, and a, b,... to conduction orbitals.

In this basis, the GKS Hamiltonian is represented as

$$H_{pq}(t) = \langle \phi_p|\hat{h}(t)|\phi_q\rangle \tag{4}$$

which is evaluated as

$$H_{pq}(t) = \langle \phi_p|\hat{h}_0 + \delta\hat{v}^\gamma + \hat{X}_{val}^\gamma + \hat{X}_{core}^\gamma|\phi_q\rangle \tag{5}$$

As mentioned, the range−separation parameter γ is non-empirically obtained by enforcing the ionization potential theorem; further details are provided in refs 17 and 40. The time-independent part of the Hamiltonian, $\hat{h}_0$, is

$$\hat{h}_0 = -\frac{1}{2}\widehat{\nabla}^2 + \hat{v}_{eN}^{NL} + \hat{v}^\gamma[n(t = 0)] \tag{6}$$

which includes the KS potential

$$\hat{v}^\gamma[n(t)] = \hat{v}_{eN}^{local} + \int \frac{n(r', t)}{|r - r'|}dr' + \hat{v}_{XC}^{SR,\gamma}[n(t)] \tag{7}$$

Here, $\hat{v}_{XC}^{SR,\gamma}[n(t)]$ is the local short-range exchange correlation potential. The next term in the Hamiltonian is the difference between the short-range KS potential at time t and at $t = 0$

$$\delta\hat{v}^{\gamma}[n(t)] = \hat{v}^{\gamma}[n(t)] - \hat{v}^{\gamma}[n(t=0)] \tag{8}$$

The core contribution to the density is always approximated as time-independent

$$n(r, t) = n_{core}(r) + n_{val}(r, t) \tag{9}$$

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

$$\langle \phi_p | \hat{X}_{val}^{\gamma} | \phi_q \rangle = -\sum_{st} P_{st}(t)(\phi_q \phi_s | u^{\gamma} | \phi_t \phi_p) \tag{10}$$

where the long-range Coulomb kernel is $u^{\gamma}(|r - r'|) = \mathrm{erf}(\gamma|r - r'|)/|r - r'|$ and the density matrix is expressed $P_{st} = \sum_i C_{si} f_i C_{ti}$ (f_i denotes orbital occupation). Assuming real orbitals, the 4-index integral tensor is in chemists notation

$$(\phi_q \phi_s | u^{\gamma} | \phi_t \phi_p) = \int \phi_q(r)\phi_s(r)u^{\gamma}(|r - r'|)\phi_t(r')\phi_p(r')\mathrm{d}r'\mathrm{d}r \tag{11}$$

To circumvent the prohibitive cost of calculating the exchange matrix elements in eq 11, the long-range Coulomb kernel is split into two parts, one treated deterministically and one that we accurately represent through a sparse-stochastic basis. Using a stochastic resolution of the identity for the Coulomb kernel in the reciprocal space (see eq 39), the exact exchange matrix elements are

$$\langle \phi_p | \hat{X}_{val}^{\gamma} | \phi_q \rangle = -\sum_{st} P_{st}(t)\langle \phi_q \phi_s | \xi \rangle \langle \xi | \phi_t \phi_p \rangle \tag{12}$$

where $\{|\xi\rangle\}$ is a small set of auxiliary vectors, as detailed in Section 2.3. Equation 12 is efficiently computed as

$$\langle \phi_p | \hat{X}_{val}^{\gamma} | \phi_q \rangle = -\sum_{\xi i} u_{p\xi i}(t)u_{q\xi i}^{*}(t) \tag{13}$$

where we only have to store the vectors (see Section 2.3)

$$u_{p\xi i}(t) = \sum_t C_{ti}(t)\langle \xi | \phi_t \phi_p \rangle \tag{14}$$

Next, the effect of the core states on the exact exchange operator, $\hat{X}_{core}^{\gamma}$, is approximated as a time-independent scissor correction

$$\hat{X}_{core}^{\gamma} = \hat{P}\mathrm{d}\varepsilon_{\gamma,HOMO} + \hat{Q}\,\mathrm{d}\varepsilon_{\gamma,LUMO} \tag{15}$$

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

$$\mathrm{d}\varepsilon_{\gamma,a} = \sum_{f \in core}(\psi_a \phi_f | u^{\gamma} | \phi_f \psi_a) \tag{16}$$

by inserting the GKS wave function, expanded in the LDA MO basis, into the TDKS equation, and using the orthogonality of the LDA MOs, an ordinary basis set TDKS equation is obtained

$$i\partial_t C(t) = H(t)C(t) \tag{17}$$

We excite the initial state with a small perturbation, with strength parameter $\Delta \approx 10^{-4}$

$$C(t = 0^+) = \mathrm{e}^{-i\Delta D}C(t = 0) \tag{18}$$

where

$$D_{pq} = \langle \phi_p | \hat{x} | \phi_q \rangle \tag{19}$$

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

$$C(t + \mathrm{d}t) \simeq \mathrm{e}^{-i\mathrm{d}tH\left(t+\frac{\mathrm{d}t}{2}\right)}C(t) \tag{20}$$

where we use a linear extrapolation approach, $H\left(t + \frac{\mathrm{d}t}{2}\right) = 1.5H(t) - 0.5H(t - \mathrm{d}t)$.

The dipole moment along $\hat{x}$ is extracted

$$\mu(t) = \sum_{qp} P_{qp}(t)D_{pq} \tag{21}$$

The absorption spectra is then generated by Fourier analysis of the time-dependent dipole moment as usual[15]

$$S(\omega) = \frac{1}{3}\mathrm{Tr}[\sigma(\omega)] \tag{22}$$

where

$$\sigma_{ii}(\omega) = \frac{4\pi\omega}{c}\mathrm{Im}[\alpha_{ii}(\omega)] \tag{23}$$

And the polarizability is calculated from $\alpha_{ii}(\omega) = \frac{\mu_i(\omega)}{\Delta}$.

2.2. $\sigma(\omega)$ via an Iterative Chebyshev Procedure. The frequencies of the TDSE solve an eigenvalue equation, known as the Casida equation.[1] For reference, see a traditional derivation including Fock exchange in ref 43. The Casida equation is formally

$$\mathcal{L}\begin{pmatrix} f^+ \\ f^- \end{pmatrix} = \hbar\omega\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} f^+ \\ f^- \end{pmatrix} \tag{24}$$

where the Liouvillian is

$$\mathcal{L} = \begin{pmatrix} A & B \\ -B & -A \end{pmatrix} \tag{25}$$

And f^+ and f^- refer to the orbitals of the linear-response perturbation. In the space of single-particle singlet excitations, the orbital Hessians A and B are

$$A(ia; jb) = (\varepsilon_a - \varepsilon_i)\delta_{ij}\delta_{ab} + 2(ia|jb) + (ia|f_{XC}|jb) \\ - (\phi_a\phi_b|u^{\gamma}|\phi_i\phi_j)$$

$$B(ia; bj) = 2(ia|bj) + (ia|f_{XC}|bj) - (\phi_a\phi_j|u^{\gamma}|\phi_i\phi_b) \tag{26}$$

With two-electron integrals

$$(ia|jb) = \int \phi_i(r)\phi_a(r)|r - r'|^{-1}\phi_j(r')\phi_b(r')\mathrm{d}r'\mathrm{d}r \tag{27}$$

So $(ia|jb) = (ia|bj)$. Similarly to the real-time approach presented in Section 2.1, the XC kernel elements are computed within an attenuated ALDA scheme to first order.[44]

To track the response of the system to an electric field in the x-direction, one uses an initial excitation spinor composed of two dipole matrix elements

$$|\chi\rangle = \begin{pmatrix} \chi_{ia}^+ \\ \chi_{ia}^- \end{pmatrix} = \begin{pmatrix} + \langle \phi_a | x | \phi_i \rangle \\ - \langle \phi_a | x | \phi_i \rangle \end{pmatrix} \tag{28}$$

Application of the Liouvillian matrix on a general vector $(f^+, f^-)^T$ is then

$$(\mathcal{L}f^+)_{ia} = (\varepsilon_a - \varepsilon_i)f_{ia}^+ + \langle\phi_a|\delta v_H|\phi_i\rangle + \langle\phi_a|\delta v_{XC}^\gamma|\phi_i\rangle$$
$$- \langle\phi_a|y_i^+\rangle - \langle\phi_a|z_i^-\rangle \tag{29}$$

And for backward transitions

$$(\mathcal{L}f^-)_{ia} = -(\varepsilon_a - \varepsilon_i)f_{ia}^- - \langle\phi_a|\delta v_H|\phi_i\rangle - \langle\phi_a|\delta v_{XC}^\gamma|\phi_i\rangle$$
$$+ \langle\phi_a|y_i^-\rangle + \langle\phi_a|z_i^+\rangle \tag{30}$$

where we defined

$$y_i^\pm(r) = \sum_{jb} f_{jb}^\pm(r) u_{ij}^\gamma(r) \tag{31}$$

$$z_i^\pm(r) = \sum_{jb} f_{jb}^\pm(r) u_{ib}^\gamma(r) \tag{32}$$

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

$$\delta n(r) = 2\sum_{jb} (f_{jb}^+ + f_{jb}^-)\phi_j(r)\phi_b(r) \tag{33}$$

The first-order change in the short-range exchange−correlation potential is analogous to the real-time approach[45]

$$\delta v_{XC}^\gamma = \frac{1}{\Delta}(v_{XC}^\gamma[n_0 + \Delta\delta n] - v_{XC}^\gamma[n_0]) \tag{34}$$

The iterative Chebyshev approach to obtain the full, non-TDA, frequency-resolved dipole−dipole correlation function $\sigma(\omega)$ is obtained as[46]

$$\sigma(\omega) \propto \omega\langle\tilde{\chi}|\delta(\mathcal{L} - \omega)|\chi\rangle \tag{35}$$

And the delta function is calculated as

$$\delta(\mathcal{L} - \omega) = \sum_{n=0}^{N_{Chebyshev}} c_n(\omega) T_n(\tilde{\mathcal{L}}) \tag{36}$$

where $T_n(\tilde{\mathcal{L}})$ is the n'th Chebyshev polynomial and $\tilde{\mathcal{L}}$ is a scaled Liouvillian with eigenvalues between -1 and 1. The optical absorption is obtained from the residues as

$$\sigma(\omega) = \sum_{n=0}^{N_{Chebyshev}} c_n(\omega) R_n \tag{37}$$

with $R_n = \langle\tilde{\chi}|T_n(\tilde{\mathcal{L}})|\chi\rangle$ and $\tilde{\chi}_{ia}^\pm = \pm\chi_{ia}^\pm$. For the Chebyshev coefficients, $c_n(\omega)$, we use simple smoothly decaying weights.[46,47]

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

2.3. Deterministic/Fragmented-Stochastic Calculation of Long-Range Exchange.

The dominant cost of the spectra calculation is computing the action of $u^\gamma(|r - r'|)$ on orbital pairs, i.e., the exchange part in TDHF (time-dependent Hartree−Fock). To reduce the cost, we use a mixed deterministic/fragmented-stochastic approach, as developed in ref [40], which is briefly overviewed below.

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

$$u_{ij}^\gamma(r) = \mathcal{F}^{-1}\{u^\gamma(k)\langle k|\phi_i\phi_j\rangle\} \tag{38}$$

where $\mathcal{F}^{-1}$ denotes an inverse FFT. To reduce the effort in eq 38, the interaction $u^\gamma(k)$ is split between long-wavelength, low-k components that are included deterministically, while the remainder high-k terms are represented through a small sparse-stochastic basis with a constant number of terms (around 1000−2000) independent of system size. To achieve this fragmentation of $u^\gamma(k)$, the following identity is introduced

$$I = \sum_{k_{low}} |k_{low}\rangle\langle k_{low}| + \sum_{k_{high}} |k_{high}\rangle\langle k_{high}| \tag{39}$$

In this basis, the Coulomb interaction is

$$u^\gamma = \sum_{k_{low}} |k_{low}\rangle u^\gamma(k_{low})\langle k_{low}| + \sum_{k_{high}} \sqrt{u^\gamma(k_{high})}\,|k_{high}\rangle$$
$$\langle k_{high}|\sqrt{u^\gamma(k_{high})} \tag{40}$$

Note that for simplicity, we present for the case where all $u^\gamma(k)$ are positive. (In practice, some elements are negative due to the use of the Martyna−Tuckerman procedure[48] to avoid grid-reflection effect, so our actual simulations use the general formulation, as detailed in ref [40].)

We introduce now a sparse-stochastic basis $\{|\alpha\rangle\}$ (with N_α members) for the k_{high} space

$$\alpha(k_{high}) = \pm\sqrt{\frac{N_{k_{high}}}{S}} A_\alpha(k_{high}) \tag{41}$$

where $N_{k_{high}}$ 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., it is one within the length-S fragment α, and vanishes outside. In principle, for the real-time segment of the paper, we could resample the random projection in the stochastic basis at each time-step to further reduce the stochastic error. However, 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

$$\langle k_{high}|\zeta\rangle = \sqrt{u^\gamma(k_{high})}\,\alpha(k_{high}) \tag{42}$$

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

$$|\xi\rangle = \{\sqrt{u^\gamma(k_{low})}\,|k_{low}\rangle\} \oplus \{|\zeta\rangle\} \tag{43}$$

And the interaction is finally a sum of separable terms

$$u^\gamma = \sum_\xi |\xi\rangle\langle\xi| \tag{44}$$

Note that the only point in our approach where there is a stochastic error is in forming the exchange-potential operator, eq 44; once a deterministic/stochastic basis is used to form an approximate but fully Hermitian representation of the exchange-potential operator, the propagation (in real-time)

Table 1. HOMO − LUMO Gaps for Naphthalene, Fullerene, Chla, and a 476 Atom Photosystem II Hexamer Dye Reaction Center (RC-PSII)[a]

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	488	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

[a]Also shown are the number of grid-points and occupied states, the maximum number of valence and conduction states, and the range-separation parameter. All energies are in eV.

or iteration (in the frequency-dependent approach) is done deterministically and will therefore be fully stable.

Our ground-state method paper, ref 40, discusses the stochastic error in more detail, and we have not detailed it here as the exchange contribution at each time-step (or Liouvillian application in the iterative Chebyshev approach) is very small, so the error is even smaller in the TDGKS than the GKS-DFT methods. We have verified that the optical spectra obtained in the manuscript are for the most part insensitive to the size of the sparse-stochastic basis N_α sampling the high-k contribution of the exchange (e.g., $N_\alpha \sim 500-5000$ yield identical spectra). The stochastic error is much smaller than that due to using a smaller M (i.e., decreasing the N_v:N_c basis size).

3. RESULTS

We first benchmark the resulting TDGKS method on naphthalene and fullerene and then study a Chla monomer and a chlorophyll hexamer reaction-center complex found at the center of Photosystem II (RC-PSII). The Chla model has a methyl acetate ligand in place of the phytyl chain. The optimized coordinates for both dye systems are taken from ref 42. For Chla, 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 metalloporphyrin ring and depend on the laser polarization direction.[49]

The linear-response calculation requires input from a prior ngH-DFT calculation. Table 1 provides the maximum 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.

For all systems, the plotted spectra are obtained from the frequency-resolved Chebyshev approach. We verified that the real-time and frequency-based approaches give identical spectra for naphthalene, fullerene, and Chla. The resolution is fixed by setting the spectral width to the number of Chebyshev terms, $\delta\mathcal{L}/N_{Chebyshev}$, constant, with an energy broadening of roughly 0.2 eV. The real-time calculations were propagated with a time-step of dt = 0.25 au for a total number of time-steps nt = 5000.

In Table 2, the optical gaps of naphthalene and fullerene are compared with those calculated with the NWChem software package.[50] All optical gaps correspond to the first peak in the absorption spectrum. The NWChem optical gaps are obtained by selecting the lowest root excitation energy with nonzero dipole oscillator strength. The LR-TDDFT NWChem calculations use an aug-cc-pvdz Gaussian basis set and the BNL exchange−correlation functional (with the same range-separation parameter γ). Good overall agreement is obtained

Table 2. Optical Gaps (First Peak in Spectrum) of Naphthalene, Fullerene, Chla, and RC-PSII Using N_v^{max}:N_c^{max} States[a]

system	optical gap	reference
naphthalene	4.68	4.52 (NWChem)
fullerene	3.33	3.30 (NWChem)
Chla	2.01, 2.20	1.99, 2.30 (Expt.)
RC-PSII	2.20	1.95 (Expt.)

[a]Q_y and Q_x peaks are given for Chla. Comparisons to available literature values are included for dye systems. All energies are in eV. Reference values for naphthalene and fullerene are from LR-TDDFT calculations with NWChem software ref 50. Chla and RC-PSII are experimental values from refs 41 and 42.

between our new method and the established software for these two hydrocarbon systems.

Figures 1 and 2 show convergence of the optical gaps of naphthalene and Chla with respect to the explicit number of
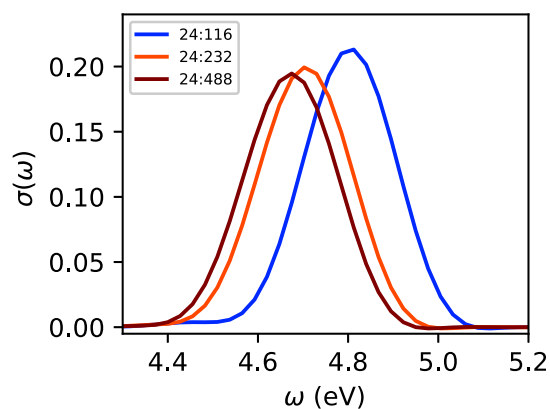

Figure 1. Convergence of the optical gap of naphthalene with respect to conduction basis N_c. The labels correspond to N_v:N_c. The frequency-domain approach spectra is used here and throughout the other figures, but we confirmed that the real-time spectra are identical.

near-gap states included in the spectral calculation. For naphthalene, fullerene, and RC-PSII, we present the averaged spectrum, while the Chla peaks are specifically chosen from the x- and y-polarized spectra.

The convergence of the naphthalene optical gap, Figure 1, depends more heavily on the inclusion of the virtual orbitals as opposed to the occupied orbitals. However, for Chla, Figure 2, convergence of the Q_y and Q_x peaks requires that all occupied states and a large number of virtual states are included. For these two systems, the optical gap peaks are both converged with the maximum number of conduction states, provided in Table 1, and a larger conduction basis did not influence the optical gap peak position. Table 3 provides the Q_y and Q_x peak
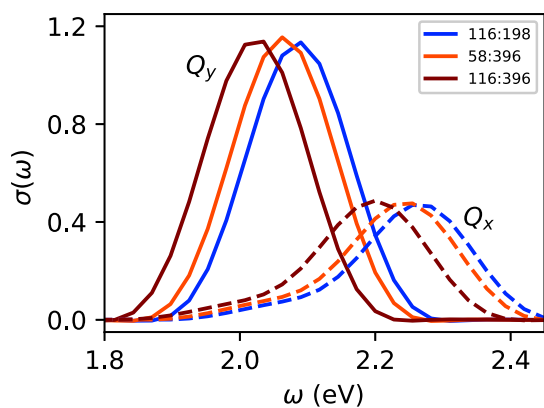
Figure 2. Convergence of the Q_y- and Q_x-optical absorption bands of Chla with active space size $N_v:N_c$.

Table 3. Convergence of the Gap between the Q_y and Q_x Bands of Chla with $N_v:N_c$[a]

$N_v:N_c$	Q_y	Q_x	$\Delta_{Q_y} - \Delta_{Q_x}$
20:40	2.23	2.45	0.22
40:80	2.17	2.34	0.17
60:120	2.12	2.31	0.19
116:198	2.09	2.26	0.17
58:396	2.06	2.26	0.20
116:396	2.01	2.20	0.19

[a]All energies are in eV.

positions as well as absolute differences for Chla with $N_v:N_c$ basis size. The distance between the peaks, $\Delta_{Q_y-Q_x}$, is well converged, within 0.03 eV, already for $N_v = 20:N_c = 40$.

Figure 3 shows that the optical gap of fullerene (C_{60}) is converged within an error of 0.03 eV for $N_v = 60:N_c = 120$ to $N_v = 120:N_c = 480$. It is known that the optical gap of fullerene is faint,[51] so to resolve the peak, we use a total of 4000 Chebyshev polynomials.

While the convergence tests above were for the desired quantity, the low-frequency spectrum, an interesting question is how the real-time dipole itself (not just the frequency-resolved spectra) converges with the basis set. Figure 3 shows this convergence for fullerene, comparing $N_v = 80:N_c = 160$ and $N_v = 120:N_c = 480$. The signal for the smaller subspace accurately captures low-frequency oscillations even though the signals do not exactly match.

Moving to the largest system, Figure 4 shows the absorption spectrum of the large dye, RC-PSII, as well as the convergence
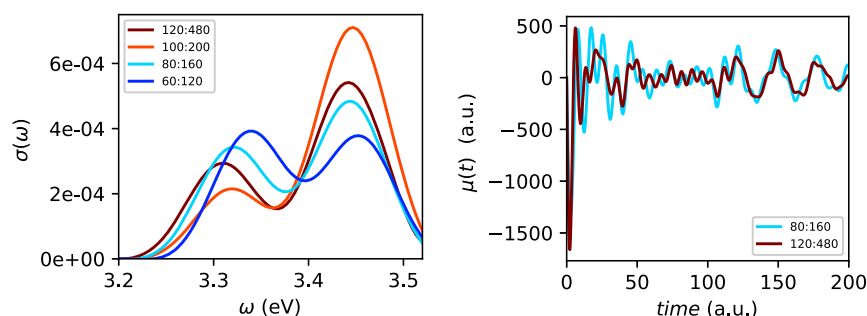

Figure 4. Absorption spectra of RC-PSII with respect to active space size $N_v:N_c$. The inset shows the convergence of optical gap with $N_v:N_c$.

of the optical gap. Unlike the smaller systems studied here, the optical gap of this much larger system converges rapidly with the fraction of valence-conduction states, with $N_v = 150:N_c = 300$ and $N_v = 200:N_c = 400$ giving nearly identical spectra.

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 only a very small stochastic error. For the four systems, we used k_{cut} = 1.8, 1.1, 0.9, 0.5 au respectively, and with the γ values provided in Table 1, the exchange kernel, $v_\gamma(k) \propto \exp(-k^2/4\gamma^2)/k^2$, is numerically tiny for the high-k stochastically sampled space.

The chosen k_{cut} values correspond to $N_{k_{low}} \simeq 5000$ low-k elements. Each system also uses $N_\alpha = 1,000$ sparse vectors for the high-k space. An overall auxiliary basis of $N_\xi \simeq 6,000$ is then used, irrespective of system size. The stochastic error in our TDDFT calculations is then negligible, much less than 0.01 eV. A more detailed breakdown of the stochastic error is given in ref 40.

4. DISCUSSION

We presented here a GKS-TDDFT formalism for optical spectra that efficiently calculates long-range exact exchange by a mixed deterministic/fragmented stochastic reciprocal-space

Figure 3. (Left) Convergence of optical gap of fullerene with respect to $N_v:N_c$. (Right) Time-dependent dipole signal comparison for fullerene using $N_v = 60:N_c = 120$ and $N_v = 120:N_c = 480$.
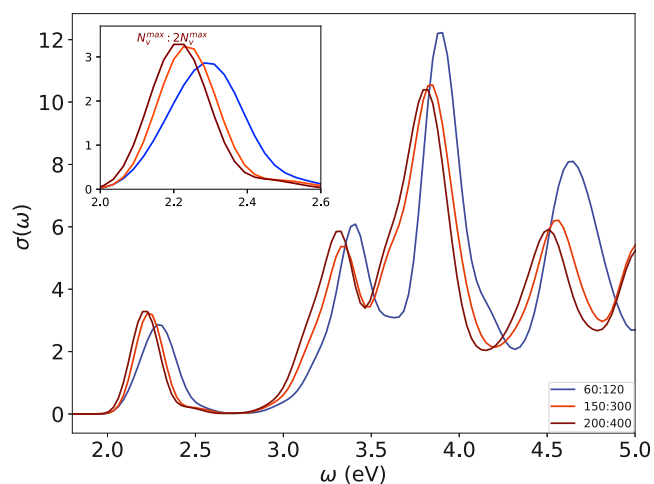
grid approach. Both a real-time and a frequency domain formulation were studied. These approaches converge the optical gap readily with the number of valence MOs, especially as the system gets larger. Thus, for the largest system, RC-PSII, it is enough to use $N_v = 150$ rather than the total number of occupied states $N_o = 660$. Additionally, the dimension of the conduction basis for the converged hexamer calculation is even smaller than the number of occupied orbitals. The number of required near-gap basis functions decreases relative to the number of occupied electrons as the system size grows due to the response becoming increasingly concentrated near the Fermi energy.

The cost of calculating the exchange matrix elements is reduced in the present approach to $\sim M^2 N_v N_\xi$, where $M = N_v + N_c$.[40] For the optical gap calculations, we use a constant number of N_ξ, and N_v and M do not increase linearly with the number of total MOs in the system. The spectral calculations are even less sensitive to the stochastic basis size N_α than are the ground-state GKS calculations. The combination of all of our techniques thus culminates in a method that is particularly well suited to extract the optical properties for systems with a large number of electrons.

To converge the optical gap in the naphthalene and Chla simulations, more conduction states are needed, compared to valence states. When $N_v < N_o$, the core correction to the exchange accounts for much of the effect of the excluded occupied states by shifting 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 working in a pair natural orbital basis to maximize the transition dipole for a cheaper ALDA-TDDFT calculation.[52]

The convergence with respect to the selected active space should be significantly improved when extending this approach to the GW-BSE method because the screened Coulomb interaction is much weaker than the exchange interaction in the current TDGKS formalism. Specifically, the inclusion of exact exchange in TDGKS yields a better description of charge-transfer excitations than in TDDFT with local or semilocal functionals; however, one needs explicit inclusion of the electron−hole attractive interaction for an even more accurate absorption spectrum. Our recent formulation of the stochastic GW-BSE recasts the expensive screened exchange term to a translationally invariant fitted exchange interaction and a small difference that is sampled stochastically.[46] This shifts the bulk of the computational effort to calculating the convolutions, as in TDHF, for which we substantially improved the efficiency here.

In addition to the extension of this LR-TDDFT approach to the BSE, there will be several other developments/extensions of the presented method.

Fragmented stochastic exchange could 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.[53] OPAW enables the 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 fragmented-stochastic GKS-TDDFT will be excited-state nuclear gradients. Previous works computing ionic forces in TDDFT were mostly limited to LDA functionals, and the presented method significantly reduces

the cost of exchange, making long-RSH functionals more accessible.[54] Note also a recent work with local hybrids.[55]

Going beyond the linear-response regime, the real-time approach would be particularly suited for calculating the hyperpolarizability of molecules as well as simulating strong-field phenomena, such as strong-field ionization and high harmonic generation. Even though most practical calculations are done on small systems, large grids are required to resolve the electron density when it is far from equilibrium. We expect to see then 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.

■ AUTHOR INFORMATION

Corresponding Author

Daniel Neuhauser − *Department of Chemistry and Biochemistry, and California Nanoscience Institute, UCLA, Los Angeles, California 90095-1569, United States;* ⊙ orcid.org/0000-0003-3160-386X; Email: dxn@ucla.edu

Authors

Mykola Sereda − *Department of Chemistry and Biochemistry, and California Nanoscience Institute, UCLA, Los Angeles, California 90095-1569, United States;* ⊙ orcid.org/0000-0002-3226-900X

Tucker Allen − *Department of Chemistry and Biochemistry, and California Nanoscience Institute, UCLA, Los Angeles, California 90095-1569, United States*

Nadine C. Bradbury − *Department of Chemistry and Biochemistry, and California Nanoscience Institute, UCLA, Los Angeles, California 90095-1569, United States;* ⊙ orcid.org/0000-0002-1214-113X

Khaled Z. Ibrahim − *Computer Science Department, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jctc.4c00260

Author Contributions

§M.S. and T.A. are co-first authors.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Casida, M. E. Time-Dependent Density Functional Response Theory for Molecules. *World Sci.* **1995**, *1*, 155−192.

(2) Jamorski, C.; Casida, M. E.; Salahub, D. R. Dynamic polarizabilities and excitation spectra from a molecular implementation of time-dependent density-functional response theory: N_2 as a case study. *J. Chem. Phys.* **1996**, *104*, 5134−5147.

(3) Yabana, K.; Bertsch, G. F. Time-Dependent Local-Density Approximation in Real Time: Application to Conjugated Molecules. *Int. J. Quantum Chem.* **1999**, *75*, 55−66.

(4) Yabana, K.; Bertsch, G. Optical response of small carbon clusters. *Z. Phys. D: At., Mol. Clusters* **1997**, *42*, 219−225.

(5) Lopata, K.; Van Kuiken, B. E.; Khalil, M.; Govind, N. Linear-Response and Real-Time Time-Dependent Density Functional Theory Studies of Core-Level Near-Edge X-Ray Absorption. *J. Chem. Theory Comput.* **2012**, *8*, 3284−3292.

(6) DeBeer George, S.; Petrenko, T.; Neese, F. Time-dependent density functional calculations of ligand K-edge X-ray absorption spectra. *Inorg. Chim. Acta* **2008**, *361*, 965−972.

(7) Lermé, J.; Palpant, B.; Prével, B.; Cottancin, E.; Pellarin, M.; Treilleux, M.; Vialle, J.; Perez, A.; Broyer, M. Optical properties of gold metal clusters: A time-dependent local-density-approximation investigation. *Eur. Phys. J. D* **1998**, *4*, 95−108.

(8) Qian, X.; Li, J.; Lin, X.; Yip, S. Time-dependent density functional theory with ultrasoft pseudopotentials: Real-time electron propagation across a molecular junction. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2006**, *73*, 035408.

(9) Baer, R.; Neuhauser, D. Ab Initio Electrical Conductance of a Molecular Wire. *Int. J. Quantum Chem.* **2003**, *91*, 524−532.

(10) Baer, R.; Gould, R. A method for ab initio nonlinear electron-density evolution. *J. Chem. Phys.* **2001**, *114*, 3385−3392.

(11) Zelovich, T.; Kronik, L.; Hod, O. State Representation Approach for Atomistic Time-Dependent Transport Calculations in Molecular Junctions. *J. Chem. Theory Comput.* **2014**, *10*, 2927−2941.

(12) Ding, F.; Van Kuiken, B. E.; Eichinger, B. E.; Li, X. An efficient method for calculating dynamical hyperpolarizabilities using real-time time-dependent density functional theory. *J. Chem. Phys.* **2013**, *138*, 064104.

(13) Parker, S. M.; Rappoport, D.; Furche, F. Quadratic Response Properties from TDDFT: Trials and Tribulations. *J. Chem. Theory Comput.* **2018**, *14*, 807−819.

(14) Perdew, J. P. Density Functional Theory and the Band Gap Problem. *Int. J. Quantum Chem.* **1985**, *28*, 497−523.

(15) Li, X.; Govind, N.; Isborn, C.; DePrince, A. E.; Lopata, K. Real-Time Time-Dependent Electronic Structure Theory. *Chem. Rev.* **2020**, *120*, 9951−9993.

(16) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648−5652.

(17) Baer, R.; Livshits, E.; Salzner, U. Tuned Range-Separated Hybrids in Density Functional Theory. *Annu. Rev. Phys. Chem.* **2010**, *61*, 85−109.

(18) Herbert, J. M.. In *Theoretical and Computational Photochemistry* García-Iriepa, C., Marazzi, M., Eds.; Elsevier, 2023; pp 69−118.

(19) Tawada, Y.; Tsuneda, T.; Yanagisawa, S.; Yanai, T.; Hirao, K. A long-range-corrected time-dependent density functional theory. *J. Chem. Phys.* **2004**, *120*, 8425−8433.

(20) Baer, R.; Neuhauser, D. Density Functional Theory with Correct Long-Range Asymptotic Behavior. *Phys. Rev. Lett.* **2005**, *94*, 043002.

(21) Leininger, T.; Stoll, H.; Werner, H.-J.; Savin, A. Combining long-range configuration interaction with short-range density functionals. *Chem. Phys. Lett.* **1997**, *275*, 151−160.

(22) Yanai, T.; Tew, D. P.; Handy, N. C. A new hybrid exchange−correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393*, 51−57.

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

(24) Lopata, K.; Reslan, R.; Kowalska, M.; Neuhauser, D.; Govind, N.; Kowalski, K. Excited-State Studies of Polyacenes: A Comparative Picture Using EOMCCSD, CR-EOMCCSD(T), Range-Separated (LR/RT)-TDDFT, TD-PM3, and TD-ZINDO. *J. Chem. Theory Comput.* **2011**, *7*, 3686−3693.

(25) Karolewski, A.; Stein, T.; Baer, R.; Kümmel, S. Communication: Tailoring the optical gap in light-harvesting molecules. *J. Chem. Phys.* **2011**, *134*, 151101.

(26) Kronik, L.; Stein, T.; Refaely-Abramson, S.; Baer, R. Excitation Gaps of Finite-Sized Systems from Optimally Tuned Range-Separated Hybrid Functionals. *J. Chem. Theory Comput.* **2012**, *8*, 1515−1531.

(27) Aquilante, F.; Pedersen, T. B.; Lindh, R. Low-cost evaluation of the exchange Fock matrix from Cholesky and density fitting representations of the electron repulsion integrals. *J. Chem. Phys.* **2007**, *126*, 194106.

(28) Manzer, S. F.; Epifanovsky, E.; Head-Gordon, M. Efficient Implementation of the Pair Atomic Resolution of the Identity Approximation for Exact Exchange for Hybrid and Range-Separated Density Functionals. *J. Chem. Theory Comput.* **2015**, *11*, 518−527.

(29) Qin, X.; Liu, J.; Hu, W.; Yang, J. Interpolative Separable Density Fitting Decomposition for Accelerating Hartree−Fock Exchange Calculations within Numerical Atomic Orbitals. *J. Phys. Chem. A* **2020**, *124*, 5664−5674.

(30) Medves, M.; Fronzoni, G.; Stener, M. Optimization of density fitting auxiliary Slater-type basis functions for time-dependent density functional theory. *J. Comput. Chem.* **2022**, *43*, 1923−1935.

(31) Sharma, S.; White, A. F.; Beylkin, G. Fast Exchange with Gaussian Basis Set Using Robust Pseudospectral Method. *J. Chem. Theory Comput.* **2022**, *18*, 7306−7320.

(32) Liu, J.; Herbert, J. An efficient and accurate approximation to time-dependent density functional theory for systems of weakly coupled monomers. *J. Chem. Phys.* **2015**, *143*, 034106.

(33) Ko, H.-Y.; Santra, B.; DiStasio, R. A. J. Enabling Large-Scale Condensed-Phase Hybrid Density Functional Theory-Based Ab Initio Molecular Dynamics II: Extensions to the Isobaric−Isoenthalpic and Isobaric−Isothermal Ensembles. *J. Chem. Theory Comput.* **2021**, *17*, 7789−7813.

(34) Hutter, J. 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.* **2003**, *118*, 3928−3934.

(35) Boffi, N. M.; Jain, M.; Natan, A. Efficient Computation of the Hartree−Fock Exchange in Real-Space with Projection Operators. *J. Chem. Theory Comput.* **2016**, *12*, 3614−3622.

(36) Liu, J.; Hu, W.; Yang, J. Accelerating Linear-Response Time-Dependent Hybrid Density Functional Theory with Low-Rank Decomposition Techniques in the Plane-Wave Basis. *J. Chem. Theory Comput.* **2022**, *18*, 6713−6721.

(37) Medves, M.; Sementa, L.; Toffoli, D.; Fronzoni, G.; Fortunelli, A.; Stener, M. An efficient hybrid scheme for time dependent density functional theory. *J. Chem. Phys.* **2020**, *152*, 184104.

(38) D'Antoni, P.; Medves, M.; Toffoli, D.; Fortunelli, A.; Stener, M.; Visscher, L. A Resolution of Identity Technique to Speed up TDDFT with Hybrid Functionals: Implementation and Application to the Magic Cluster Series $Au_{8n+4}(SC_6H_5)_{4n+8}$ (n = 3−6). *J. Phys. Chem. A* **2023**, *127*, 9244−9257.

(39) Zhang, X.; Lu, G.; Baer, R.; Rabani, E.; Neuhauser, D. Linear-Response Time-Dependent Density Functional Theory with Stochastic Range-Separated Hybrids. *J. Chem. Theory Comput.* **2020**, *16*, 1064−1072.

(40) Bradbury, N. C.; Allen, T.; Nguyen, M.; Neuhauser, D. Deterministic/Fragmented-Stochastic Exchange for Large-Scale Hybrid DFT Calculations. *J. Chem. Theory Comput.* **2023**, *19*, 9239−9247.

(41) Gruber, E.; Kjær, C.; Nielsen, S. B.; Andersen, L. H. Intrinsic Photophysics of Light-harvesting Charge-tagged Chlorophyll a and b Pigments. *Chem.—Eur. J.* **2019**, *25*, 9153−9158.

(42) Förster, A.; Visscher, L. Quasiparticle Self-Consistent GW-Bethe-Salpeter Equation Calculations for Large Chromophoric Systems. *J. Chem. Theory Comput.* **2022**, *18*, 6779−6793.

(43) Negele, J. W. The mean-field theory of nuclear structure and dynamics. *Rev. Mod. Phys.* **1982**, *54*, 913−1015.

(44) Baer, R.; Neuhauser, D. Real-time linear response for time-dependent density-functional theory. *J. Chem. Phys.* **2004**, *121*, 9803−9807.

(45) Baroni, S.; Giannozzi, P.; Testa, A. Green's-Function Approach to Linear Response in Solids. *Phys. Rev. Lett.* **1987**, *58*, 1861−1864.

(46) Bradbury, N. C.; Allen, T.; Nguyen, M.; Ibrahim, K. Z.; Neuhauser, D. Optimized attenuated interaction: Enabling stochastic Bethe−Salpeter spectra for large systems. *J. Chem. Phys.* **2023**, *158*, 154104.

(47) Weiße, A.; Wellein, G.; Alvermann, A.; Fehske, H. The kernel polynomial method. *Rev. Mod. Phys.* **2006**, *78*, 275−306.

(48) Martyna, G. J.; Tuckerman, M. E. A reciprocal space based method for treating long range interactions in ab initio and force-field-based calculations in clusters. *J. Chem. Phys.* **1999**, *110*, 2810−2821.

(49) Gouterman, M. Spectra of porphyrins. *J. Mol. Spectrosc.* **1961**, *6*, 138−163.

(50) Aprà, E.; Bylaska, E. J.; de Jong, W. A.; Govind, N.; Kowalski, K.; Straatsma, T. P.; Valiev, M.; van Dam, H. J. J.; Alexeev, Y.; Anchell, J.; et al. NWChem: Past, present, and future. *J. Chem. Phys.* **2020**, *152*, 184102.

(51) Bauernschmitt, R.; Ahlrichs, R.; Hennrich, F. H.; Kappes, M. M. Experiment versus Time Dependent Density Functional Theory Prediction of Fullerene Electronic Absorption. *J. Am. Chem. Soc.* **1998**, *120*, 5052−5059.

(52) Martin, R. L. Natural transition orbitals. *J. Chem. Phys.* **2003**, *118*, 4775−4777.

(53) Nguyen, M.; Duong, T.; Neuhauser, D. Time-dependent density functional theory with the orthogonal projector augmented wave method. *J. Chem. Phys.* **2024**, *160*, 144101.

(54) Zhang, X.; Lu, G. Subspace formulation of time-dependent density functional theory for large-scale calculations. *J. Chem. Phys.* **2015**, *143* (6), 064110.

(55) Grotjahn, R.; Furche, F.; Kaupp, M. Development and Implementation of Excited-State Gradients for Local Hybrid Functionals. *J. Chem. Theory Comput.* **2019**, *15*, 5508−5522.

(56) Boerner, T. J.; Deems, S.; Furlani, T. R.; Knuth, S. L.; Towns, J. *Practice and Experience in Advanced Research Computing*; ACM, 2023, pp 173−176.