The Journal of Chemical Physics

Optimized attenuated interaction: Enabling stochastic Bethe–Salpeter spectra for large systems

Cite as: J. Chem. Phys. **158**, 154104 (2023); https://doi.org/10.1063/5.0146555
Submitted: 14 February 2023 • Accepted: 03 April 2023 • Published Online: 17 April 2023

Nadine C. Bradbury, Tucker Allen, Minh Nguyen, et al.

AIP Publishing

The Journal of Chemical Physics ARTICLE scitation.org/journal/jcp

Optimized attenuated interaction: Enabling stochastic Bethe–Salpeter spectra for large systems

Cite as: J. Chem. Phys. 158, 154104 (2023); doi: 10.1063/5.0146555
Submitted: 14 February 2023 • Accepted: 3 April 2023 •
Published Online: 17 April 2023

Nadine C. Bradbury,[1,a)] Tucker Allen,[1] Minh Nguyen,[1] Khaled Z. Ibrahim,[2] and Daniel Neuhauser[3]

AFFILIATIONS

[1] Department of Chemistry and Biochemistry, UCLA, Los Angeles, California 90095-1569, USA
[2] Computer Science Department, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, California 94720, USA
[3] Department of Chemistry and Biochemistry, and California Nanoscience Institute, UCLA, Los Angeles, California 90095-1569, USA

[a)] **Author to whom correspondence should be addressed:** nadinebradbury@ucla.edu

ABSTRACT

We develop an improved stochastic formalism for the Bethe–Salpeter equation (BSE), based on an exact separation of the effective-interaction W into two parts, $W = (W - v_W) + v_W$, where the latter is formally any translationally invariant interaction, $v_W(r - r')$. When optimizing the fit of the exchange kernel v_W to W, using a stochastic sampling W, the difference $W - v_W$ becomes quite small. Then, in the main BSE routine, this small difference is stochastically sampled. The number of stochastic samples needed for an accurate spectrum is then largely independent of system size. While the method is formally cubic in scaling, the scaling prefactor is small due to the constant number of stochastic orbitals needed for sampling W.

I. INTRODUCTION

The Bethe–Salpeter equation (BSE), a many-body perturbation theory method, is becoming increasingly popular for predicting optical spectra of chemical systems.[1] Physically, BSE goes beyond the time-dependent density functional theory (TDDFT) by the inclusion of the correct long- range exchange kernel in the effective interaction W. Numerically, however, the BSE is quite expensive, mostly due to the cost of generating the two-electron integrals of the effective interaction W, which scales formally as $O(N^4)$, or in specially optimized cases, these integrals can be made at an $O(N^2)$–$O(N^3)$ cost,[2,3] where N is the number of electrons. Due to the steep scaling, the BSE is typically applied for systems with up to about 100 valence and conduction states. However, thanks to the many advancements in the algorithms used,[4–7] the method was recently applied to a system of nearly 2000 total electrons.[8]

Recently, we developed a numerically efficient approach to the BSE that relies on a stochastic evaluation of W.[9] In the BSE, in general, W is applied on many-pair densities of states (see later for details). Since its evaluation is linear in system size, systems with hundreds of active electrons become feasible.

In this work, we go a step beyond and show that not only is the action of W obtained efficiently with a stochastic approach, but, equally important, the explicit matrix elements can be replaced by a stochastic sampling of the sea of pair-densities. This, in principle, limits the major cost of the BSE to quadratic scaling, thereby opening the possibility of calculating spectra for very large systems.

The key in our proposed approach is the numerically exact rewriting of the action of W by subtracting and adding a simple Coulomb-like interaction v_W. Thus, the stochastic sampling only needs to be applied on this small difference $W - v_W$, with the bulk of the action of W done by v_W. This stabilizes the stochastic approach, ensuring that for larger systems, we do not need more stochastic samples to represent W.

Choosing an analytical Coulombic-like interaction to substitute for W has been done before in some efficient implementations of the BSE,[10,11] but, here, we choose an optimized v_W, which is fitted to the actual W of each system. The use of v_W is also reminiscent

of TDDFT-based approaches with long-range exchange and a polarizable medium that mimics the dielectric function.[12] In our work, since the v_W is built from W, the *ab initio* nature of the BSE is retained, while still reducing the complexity of the exchange to be similar to the traditional Fock exchange.

Using v_W by itself also gives fairly reasonable spectral results. Thus, our work is not only a numerically more efficient way to calculate the BSE spectra, but gives an alternative, fairly cheap algorithm, at the same cost as the time-dependent Hartree–Fock (TDHF), which itself can be done cheaply with a stochastic approach,[11,13,14] that has an improving accuracy for increasingly large systems.

The paper is organized as follows: The methodology is reviewed in Sec. II. Section III Shows the results for a variety of medium to large carbon- based systems. Conclusions follow in Sec. IV.

II. METHODS

A. Iterative BSE formulation

We first overview the methodology for obtaining spectra from the BSE for a given W using an iterative method. The full derivation for this method from the time-dependent Hartree–Fock (TDHF) formalism was given in previous work on the BSE,[9] and good reviews of this kind of derivation can be found in Refs. 15 and 16.

The starting point is a closed shell system with $2N_{\rm occ}$ electrons. The exciton (electron–hole) basis is a set of n_o occupied (valence) states, $\phi_i, \phi_j, \ldots$, times a set of n_c conduction states, $\phi_a, \phi_b, \ldots$, which are eigenstates of a zero order (typically DFT) Hamiltonian. Furthermore, we use the Tamm–Dancoff approximation (TDA), although the approach is generalizable to the full BSE.

The starting optically excited vector f^0 corresponds to the infinitesimal change to a ground state orbital, perturbed along a coordinate, in the direction of the laser polarization, which takes the form

$$f^0_{ja} = \langle \phi_a | x | \phi_j \rangle. \tag{1}$$

The spectrum is then obtained from a matrix element $\delta(w - A)$, where A is the Liouvillian operator of the BSE matrix, which governs the motion of the excitons:

$$\sigma(\omega) \propto \omega \langle f^0 | \delta(A - \omega) | f^0 \rangle, \tag{2}$$

where the broadened delta function is obtained by a Chebyshev series,

$$\delta(A - \omega)|f^0\rangle = \sum_n c_n(\omega)|f^n\rangle, \tag{3}$$

where c_n is the numerical coefficient and f^n the Chebyshev vector, obtained by iteratively applying A on f^0. In practice, we find that the best results are obtained by simple smoothly decaying weights, in the spirit of those used in Ref. 17:

$$c_n(\omega) = \left|\frac{d\theta_\omega}{\omega}\right| \cos^2\left(\frac{\pi n}{2N_{\rm cheby}}\right) \cos(n\theta_\omega), \tag{4}$$

where $N_{\rm cheby}$ is the number of Chebyshev terms used, which determines the frequency resolution. Here, we introduced the Chebyshev angle $\theta_\omega \equiv \cos^{-1}(\omega/\delta A)$, where δA is an upper bound on the half-width of the spectrum of A. Note that without the $|d\theta_\omega/\omega|$ term, these weights would yield a delta function in θ_ω, and this term converts the overall function to a delta function over ω.

Formally, A is made from three terms: the diagonal, the Hartree, and the so-called direct terms (in a somewhat confusing notation, since it resembles Fock exchange):

$$A_{ia,jb} = (\varepsilon_a - \varepsilon_i + \Delta)\delta_{ij}\delta_{jb} + \kappa(ia|jb) - (\phi_a\phi_b|W|\phi_i\phi_j), \tag{5}$$

where we introduced the electron and hole energies associated with the respective zero order eigenstates as ε_a and ε_i, respectively, while the round brackets refer to an $(r, r|\cdots|r', r')$ notation, and $\kappa = 2$ is used for singlet excitations and $\kappa = 0$ for triplet excitations. Δ is a scissor shift that corrects the gap to match the accurate GW calculations and could, if wished, depend on the exciton (i, a) indices—as is especially important for small systems.[18,19] In practice, we use the cheap sGW, i.e., stochastic GW (see below), to calculate the scissor term,[20,21] and for further accuracy, we implement the scissor-shift, self-consistent GW_0 approach, labeled ΔGW_0,[22] which post-processes the results of sGW and generally raises the gap by a few tenths of eV.

The Hartree integral is (assuming real orbitals)

$$(ia|jb) = \int \phi_i(r)\phi_a(r)v(r - r')\phi_j(r')\phi_b(r')drdr', \tag{6}$$

where $v(r - r') = 1/|r - r'|$ is the Coulomb interaction. Finally, the most numerically costly part involves the effective interaction

$$(\phi_a\phi_b|W|\phi_i\phi_j) = \int \phi_a(r)\phi_b(r)W(r, r')\phi_i(r')\phi_j(r')drdr'. \tag{7}$$

Note that W refers to the static part of the effective interaction, and we ignore here the effects of the dynamic part.

Numerically, one acts with A on an arbitrary vector f as follows:

$$g_{ia} \equiv (Af)_{ia} = (\varepsilon_a - \varepsilon_i + \Delta)f_{ia} + \frac{\kappa}{2}\langle \phi_a | \delta v_H | f_i \rangle - \langle \phi_a | y_i \rangle, \tag{8}$$

where the grid-representation of the exciton is

$$f_i(r) = \sum_b f_{ib}\phi_b(r), \tag{9}$$

while the exciton Coulomb density, $\delta n(r) = 4\sum_j f_j(r)\phi_j(r)$, is used to generate the Hartree potential

$$\delta v_H(r) = \int \frac{\delta n(r')}{r - r'} dr'. \tag{10}$$

The numerically expensive part in Eq. (8) comes from the direct term, involving the action of the effective interaction

$$y_i(r) \equiv \sum_j W_{ij}(r)f_j(r), \tag{11}$$

where

$$W_{ij}(r) \equiv \int W(r, r')\phi_i(r')\phi_j(r')dr'. \tag{12}$$

In our recent work,[9] we used the stochastic, time-dependent Hartree (i.e., stochastic W) approach,[14] developed originally for

sGW,[20,21] to evaluate each specific W_{ij} function in linear scaling; see Ref. 9 for full details on this step of the method. The application of stochastic W makes it feasible to study systems with up to several hundred valence states. Nevertheless, as there are $\simeq N_v^2/2$ such terms for N_v valence states, the overall cost is cubic in system size with a large pre-factor, so that including more than ≈300 valence states will be numerically challenging.

Finally, we note that this method directly obtains the BSE spectrum without capturing all excitonic eigenstates of the system. As demonstrated in our previous work on the subject,[9] specific excitonic states (for instance but not limited to the lowest energy exciton), can be purified out from this formalism using the filter diagonaliation approach.[23,24]

B. Stochastic evaluation of matrix elements

To overcome the scaling problem, we use a stochastic representation of the sum. Specifically, we define a stochastic process, made from "instances." For each such instance, we define two independent stochastic vectors,

$$\begin{aligned}\tilde{\beta}(r) &= \sum_l \tilde{\beta}_l \phi_l(r),\\ \tilde{\tilde{\beta}}(r) &= \sum_l \tilde{\tilde{\beta}}_l \phi_l(r),\end{aligned} \tag{13}$$

where $\tilde{\beta}_l = \pm 1$, $\tilde{\tilde{\beta}}_l = \pm 1$.

Using

$$\left\{\tilde{\beta}_i \tilde{\tilde{\beta}}_j\right\} = \left\{\tilde{\beta}_i \tilde{\tilde{\beta}}_j\right\} = \delta_{ij}, \tag{14}$$

where curly brackets denote an average over many stochastic instances, consequently

$$\left\{\tilde{\beta}_i \tilde{\tilde{\beta}}_j \beta(r)\right\} = \phi_i(r)\phi_j(r), \tag{15}$$

where

$$\beta(r) \equiv \tilde{\beta}(r)\tilde{\tilde{\beta}}(r). \tag{16}$$

Inserting the relations above to the numerically expensive effective potential term in Eq. (8), the latter becomes

$$y_i(r) = \left\{\tilde{\beta}_i \langle r|W|\beta\rangle f_{\tilde{\tilde{\beta}}}(r)\right\}, \tag{17}$$

where we defined

$$f_{\tilde{\tilde{\beta}}}(r) = \sum_j \tilde{\tilde{\beta}}_j f_j(r), \tag{18}$$

while $\langle r|W|\tilde{\beta}\tilde{\tilde{\beta}}\rangle \equiv \int W(r,r')\tilde{\beta}(r)\tilde{\tilde{\beta}}(r)dr'$.

The resulting algorithm is thus quite simple. A large but finite number of stochastic instances, N_β, is defined. Then, one applies W (calculated itself stochastically) on the stochastic representation of the valence density $\tilde{\beta}(r)\tilde{\tilde{\beta}}(r)$, to yield a set of N_β vectors, $\langle r|W|\tilde{\beta}\tilde{\tilde{\beta}}\rangle$, which is stored and used in the Chebyshev iterative step, $f \to Af$. The formulas are further detailed in Sec. II C.

C. Optimized attenuated interaction

1. Sampling a small difference

The formalism above is clearly a member of our stochastic approaches to quantum chemistry.[11,13,20,25,26] The key in these approaches is the replacement of individual molecular orbitals by random orbitals, that are stochastic combination of individual orbitals. For example, a valence orbital is replaced by a stochastic combination of valence orbitals, etc.

A key practical point in this paradigm is that it is best to stochastically sample numerically small quantities. This is best achieved by sampling just the difference between the desired quantity and a simpler one, i.e., writing

$$W = \{W - v_W\} + v_W, \tag{19}$$

where curly brackets indicate again a statistical average and $v_W(r,r')$ is an interaction that is "cheap" to act with. Here, we use the simplest such form, a translationally invariant, two-body interaction,

$$v_W(r,r') = v_W(r-r'). \tag{20}$$

The specifics of v_W are delineated later.

Using this decomposition, the action of W, Eq. (17), is modified to

$$\begin{aligned}y_i(r) &= \left\{\tilde{\beta}_i \langle r|W - v_W|\beta\rangle f_{\tilde{\tilde{\beta}}}(r)\right\} + \sum_j f_j(r)\langle r|v_W|\phi_i\phi_j\rangle\\ &= \left\{\tilde{\beta}_i \langle r|W - v_W|\beta\rangle f_{\tilde{\tilde{\beta}}}(r)\right\} + \sum_j f_j(r) v_{W,ij}(r),\end{aligned} \tag{21}$$

where $v_{W,ij}(r) = \int v_W(r,r')\phi_i(r')\phi_j(r')dr'$.

2. Optimizing the effective interaction potential

The equation above is exact, no matter what v_W is—a better choice of v_W would simply lead to faster convergence of the sampling of $\{W - v_W\}$. Furthermore, to avoid the singularities of the Coulomb potential, we fit only the polarization part, i.e., $W - v_W = W_{\rm pol} - v_{W\rm pol}$, where $W_{\rm pol} = W - v(k)$, and, similarly, for $v_{W\,\rm pol}$. Here, $v(k)$ is the Coulomb potential for finite systems, which is obtained with the Martyna–Tuckerman approach;[27] this potential is the usual $4\pi/k^2$ at high momenta, but levels off to a finite large value at $k = 0$.

Given an arbitrary large system, we can ask what will be the optimized $v_W(r-r')$. Interestingly, the stochastic paradigm answers that question easily. Specifically, optimize the functional

$$J = \sum_{ij} (\phi_i\phi_j|(W - v_W)^2|\phi_i\phi_j), \tag{22}$$

where again i,j are the occupied states. Calculate the sum, then, stochastically:

$$J = \left\{\langle\beta|(W - v_W)^2|\beta\rangle\right\} = \left\{\int |\langle k|W - v_W|\beta\rangle|^2\, dk\right\}, \tag{23}$$

where, as before, β is a stochastic combination of the occupied two-electron product terms from Eq. (16).

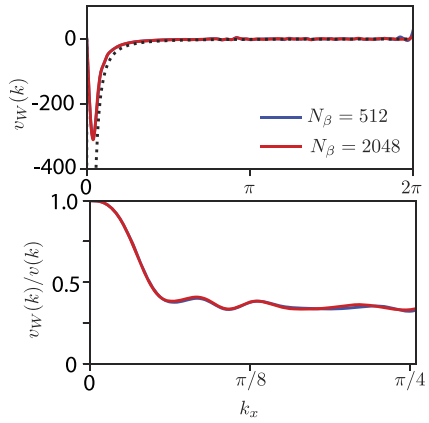

FIG. 1. (Top): The fitted $v_{W\text{pol}}(k)$ potentials (red) and the bare Coulomb interaction $v(k)$ (black dots) for $C_{96}H_{24}$, shown for a range of k_x, for $k_y = k_z = 0$. (Bottom): The ratio $v_W(k)/v(k)$ for this system, calculated for the same k-value range as in the top panel. The results converge quickly with the number of stochastic sampling functions, N_β.

Since our choice of v_W is diagonal in momentum space, $\langle k|v_W|\beta\rangle = v_W(k)\beta(k)$, it is easy to show that the optimized J, giving $\delta J/\delta v_W^*(k) = 0$, is obtained with

$$v_W(k) = \frac{\{\beta^*(k)\langle k|W|\beta\rangle\}}{\{|\langle k|\beta\rangle|^2\}}. \tag{24}$$

In practice, a very small numbers of terms, typically, $N_\beta \approx 500$, is sufficient to converge the values of $v_W(k)$. This convergence is demonstrated in Fig. 1.

3. Replacing W by the optimized attenuated interaction

If v_W is a good enough approximation to W so that the objective J is sufficiently small, we may even, as mentioned, throw out the stochastic $\{W - v_W\}$ term in Eq. (21), i.e., approximate

$$y_i(r) \simeq \sum_j f_j(r) v_{W,ij}(r). \tag{25}$$

More generally, we can approximate the full BSE by replacing W by v_W, converting, thereby, the equation to the TDHF-like with a modified Fock kernel, where $|r - r'|^{-1}$ is replaced by v_W. Note that simplified forms have been used to approximate W, see, e.g., Ref. 10, but, here, the optimized attenuated interaction is based on the true system-dependent $W(r, r')$, yielding a fully *ab initio* approach. We label the resulting method as Time-Dependent Optimized-Attenuated-Interaction (TDOAI).

D. Overall algorithm

The overall algorithm is then

- First, a set of stochastic-GW calculations on the HOMO and LUMO is performed to find the necessary scissor shift.
- Second, a set of N_β random representations, β, of the occupied-states product is calculated and stored, as per Eqs. (13) and (16).
- Each of these βs is then used as input for a stochastic-GW calculation, yielding the action of the static effective interaction $\langle r|W|\beta\rangle$.
- The Fourier-components of the optimized attenuated interaction, $v_W(k)$, are then calculated from Eq. (24).

At that point, one has the optimized attenuated interaction, but there are still several possibilities for the dynamics, i.e., how to propagate and solve the BSE and with which terms included. We summarize four such possibilities, and Sec. III below exemplifies the first two, which use the Chebyshev approach [based on Eqs. (2), (3), and (8)].

1. The BSE kernel, in the Tamm–Dancoff approximation, can be calculated by stochastically sampling the $\{W - v_W\}$ difference, Eq. (21).
2. Another direction is to ignore the $W - v_W$ term and act only with the optimized effective interaction, v_W, i.e., use Eq. (25) instead of Eq. (21).
3. One could use the first option, but go past the Tamm–Dancoff approximation, i.e., include off-diagonal terms. The simplest option, without increasing the numerical effort substantially, would be to use the optimized attenuated interaction in the off-diagonal portion of the BSE. Since the off-diagonal BSE term (i.e., the term that goes beyond the Tamm–Dancoff approximation) is quite small, it should be accurate to replace it in W entirely by v_W. This would reduce the numerical cost substantially compared to the full cost of applying $\{W - v_W\}$ stochastically in the off-diagonal term, which would have required a different sampling of the action of W, this time acting on an occupied–unoccupied pair density.
4. Finally, just like the second option above, we could use only the attenuated interaction, while avoiding the Tamm–Dancoff approximation. This could be done by either extending the exciton vector space to go beyond the TDA, or by replacing the Chebyshev method altogether by a full-fledged TDHF-like study that uses stochastic-exchange,[11] but, here, the optimized TDOAI exchange-interaction v_W would be employed; this direction would be pursued in a latter publication.

III. RESULTS

We demonstrate the new method on a sample set of hydrocarbons, including linear acenes, polycyclic-aromatic hydrocarbons (PCH), and fullerene- based systems; see Table I and Fig. 2. The structures for these molecules were taken from Refs. 28 and 29, and the open source library associated with Ref. 30. For all systems, we use a generous box size, extended at least 6 bohrs beyond the edge of the molecule, with a grid spacing of 0.5 bohr. For the planar molecules, we use a grid size of 15 bohrs in the out-of-plane direction such that no size effects are seen on the DFT bandgap. All DFT calculations were performed with norm-conserving pseudo-potentials and used the PW-MT (plane wave Martyna-Tuckerman) local-density approximation (LDA) exchange-correlation functional.[14,31,32]

TABLE I. The grid size, number of occupied orbitals, and the chosen number of valence and conduction subset sizes for each system. The final column shows the remaining fraction of the polarization interaction, W_{pol}, not captured by the optimally fitted interaction, $v_{W\text{pol}}$.

System	N_g	N_o	N_v	N_c	$\frac{\langle (W_{\text{pol}} - v_{W\,\text{pol}})^2 \rangle}{\langle W_{\text{pol}}^2 \rangle}$
Nap	50 688	24	16	40	0.18
Tet	76 800	42	24	64	0.13
Hex	113 520	60	36	80	0.11
Oct	132 000	78	45	100	0.10
Cor	69 984	45	27	70	0.18
C_{60}	195 112	120	64	120	0.25
Kek	147 000	108	54	110	0.09
$C_{96}H_{24}$	324 480	204	100	500	0.09
10-CPP + C_{60}	381 024	260	100	500	0.13

To determine the correct scissor shift, Δ, in Eq. (8), we first correct the DFT bandgap through a stochastic GW calculation,[20,21,33] done self-consistently.[22] Furthermore, the dielectric correction, $W(k \to 0)$, was determined by a linear fit of the BSE spectra at different grid sizes.[9,34,35] The final scissor shift is then the sum of the dielectric correction and the GW bandgap correction.

The calculations of the action of W on either deterministic or stochastic DFT orbital pairs, $\langle r|W|\phi_i\phi_j\rangle$ and $\langle r|W|\beta\rangle$, respectively, were done with only ten stochastic, time-dependent orbitals. Refer to Ref. 9 for explicit details of this step. The sGW calculations were done with a broadening of 0.1 hartree. A time-step dt = 0.1 a.u. was used for a split-operator propagation, and "cleaning" (i.e., projection of the excited component of the orbitals to be orthogonal to the occupied space;see Ref. 9) was done every ten steps.

The number of samples needed for a deterministic calculation is $N_v(N_v+1)/2$ for N_v valence orbitals. In calculations where W is acting on stochastic orbitals, N_β = 2000 was generally used.

We first discuss the convergence of the fitted v_W and how it compares with W.

Figure 1 shows the convergence of the fitting of the polarization portion of v_W. The top sub figure shows $v_{W\text{pol}}$ in comparison to (minus) the bare Coulomb potential $v(k)$ for this finite system. Note that $v_{W\text{pol}}$ is automatically zero at k = 0 as W_{pol} emanates from a polarization χ_{pol}, which vanishes at k = 0 due to the orthogonality of the particle–hole pairs that make it. (For periodic systems, this effect is counteracted by the singularity of the Coulomb potential at k = 0, unlike finite systems, where $v(k=0)$ is large, but does not diverge, so, $W_{\text{pol}}(k=0)$ vanishes.)

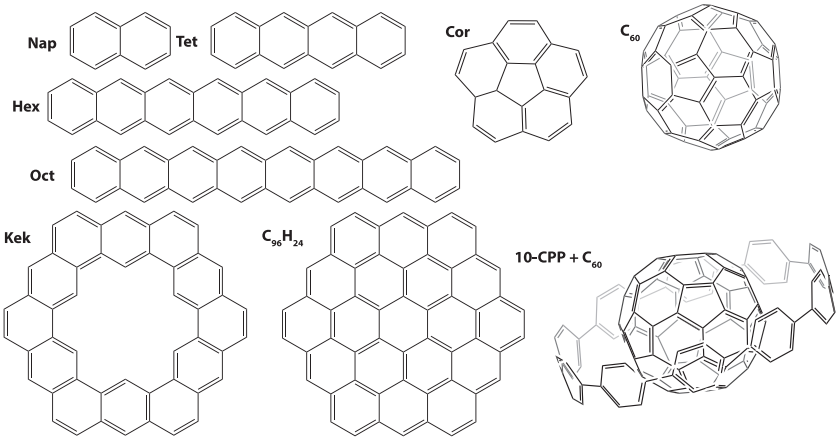

FIG. 2. Structures and abbreviations for all the systems used in this paper.

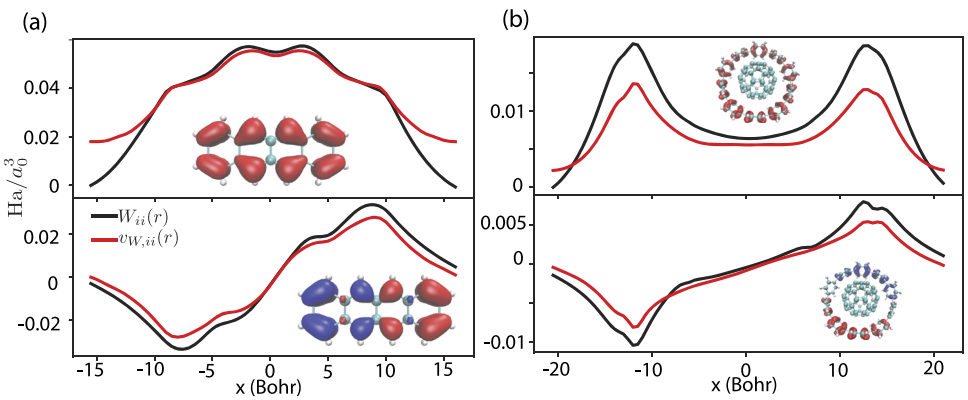

FIG. 3. X-axis slice of $\langle r|W|\phi_i\phi_j\rangle$ and $\langle r|v_W|\phi_i\phi_j\rangle$, i.e., the action of the true W (black) and the optimized v_W (red) on a two-orbital pair density. In the top row, W and v_W act on the HOMO density ($i = j$ = HOMO). In the bottom row, they act on the pair density of the HOMO × HOMO − 1 orbitals. The left part, column (a), is for tetracene, while column (b) shows the same plots for the much larger 10-CPP + C_{60}.

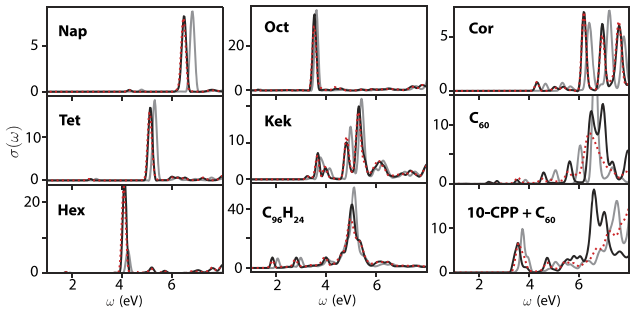

FIG. 4. Spectra of singlet excitations for all systems using a BSE with deterministic orbitals (black) and the fully stochastic $v_W + \{W - v_W\}$ approach using $N_\beta = 2000$ (red dots). The TDOAI calculation, where v_W is used for exchange alone, is shown in gray. For almost all cases, the stochastic approach matches the deterministic optical gap to within 0.02 eV or better; the one exception was fullerene, where $N_\beta = 5000$ was needed for convergence to 0.08 eV, in line with the lower quality of the v_W fit (Table I).

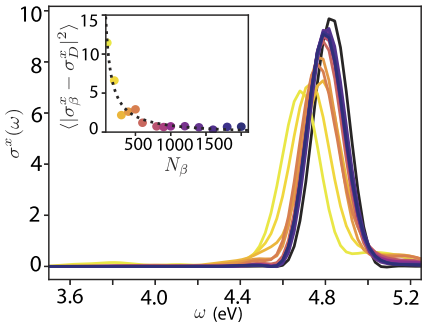

FIG. 5. X-polarization spectra for tetracene (zooming in on the dominant spectral peak at 4.75 eV) at varying levels of stochastic approximation (colors), converging to the deterministic BSE (black dashes). The inlay shows the average variance over the 0–6 eV spectral region from the deterministic spectra for each level of approximation, N_β, with the corresponding colors. The dotted curve is a fit to $1/N_\beta$.

In the bottom panel of Fig. 1, we compare the ratio between $v_W(k)$ and $v(k)$. The ratio is 1 for low k due to the finite size of the systems, but levels down at higher k values.

Table I shows the fraction of W_{pol} left for stochastic sampling after removal of $v_{W\text{pol}}$. This fraction is quite small and is clearly independent of system size. Additionally, it does not appear to change with the approximate dimensionality of the system—linear, planar, or spherical. Similarly, Fig. 3 shows, for a slice along the x-axis, the action of both W and v_W on two pair densities. The results are very similar, but the total magnitude is often decreased when applying v_W.

We now turn to the spectra. We used an iterative BSE Chebyshev procedure, and for all systems, the upper bound on the halfwidth of the Liouvillian was taken as $\delta A = 16.5$ eV. We used $N_{\text{cheby}} = 500$ terms [Eq. (4)], which is approximately equivalent to a Gaussian energy broadening with halfwidth of 0.08 eV. The effect of the broadening is negligible for the larger systems, where the spectrum is naturally quite broadened.

In Fig. 4, we show the spectra of all nine systems using a deterministic BSE,[9] the v_W only TD-OAI, and the stochastic $v_W + \{W - v_w\}$ approach of this paper. Using v_W by itself is only qualitatively accurate, but stochastic sampling of $\{W - v_W\}$ quickly restores the accuracy of the deterministic calculation.

As is clear from Fig. 5, at least $N_\beta = 300$–400 stochastic samples are needed to get a 0.1 eV accuracy on the optical gap. Generally, the low-energy spectral peaks in Fig. 4 are converged to 0.02 eV at low energies by 2000 stochastic samples. (The one exception is fullerene, where the lowest-energy spectral peak converges to only 0.08 eV at $N_\beta = 5000$; this is in line with the lower quality of the v_W fit to W for fullerene, see Table I.) Depending on the quality of v_W, we find, on average, that the $W - v_W$ formalism converges at least 20 times faster compared to that using a stochastic W by itself, i.e., without using v_W.

The rapid convergence with N_β implies that the stochastically sampled $\{W - v_W\}$ is generally numerically superior to the deterministic approach for systems with more than ≈70 calculated valence orbitals. This is because of the $N_v^2/2$ scaling of the number of pairs W_{ij} when using directly the deterministic approach, Eq. (11).

TABLE II. Gaps (eV) from stochastic DFT at the LDA level, stochastic G_0W_0, self-consistent ΔGW_0,[22] the stochastic BSE optical gap (this work), and a reference experimental optical gap.

	DFT	G_0W_0	ΔGW_0	BSE	Experimental optical gap	
Nap	3.4	7.6	8.0	4.3	4.1	36 and 37
Tet	1.6	5.1	5.4	2.7	2.6	36 and 37
Hex	0.8	3.7	3.9	1.8	1.9	38–40
Oct	0.4	2.9	3.1	1.3	1.5	39–41
Cor	3.2	6.7	7.1	4.3	3.7	42
C_{60}	1.7	4.4	4.7	2.3	1.8	43 and 44
Kek	2.1	4.8	5.1	3.2		
$C_{96}H_{24}$	1.2	3.0	3.1	1.9	2.0	45
10-CPP + C_{60}	0.7	3.3	3.5	3.5	3.4[a]	46

[a]Stabilized system complex.

In Table II, we summarize the evolution of the gap for each system. The results are in fair agreement with the experimental values, considering the Tamm–Dancoff approximation and the lack of dynamic corrections.

IV. CONCLUSIONS

We introduced here, an optimized effective potential, v_W. to reduce the magnitude of the W term in the BSE, enabling an efficient stochastic evaluation. With the introduction of v_W, the required number of stochastic orbitals is small relative to system size, thereby reducing the scaling of the method so that large system sizes are now feasible. The new algorithm was checked successfully on nine molecules of varying dimensions and sizes.

The present work overcomes the cost of the most expensive part in the BSE algorithm, preparing the action of W on the product states, by dividing W by an exchange-type potential, v_W, and a stochastically sampled remainder, $\{W - v_W\}$. There is however, lot of room for further scaling improvements. Currently, we do not implement the exchange in a particularly efficient way, so, the scaling is still cubic, but a fully stochastic exchange could further reduce the scaling.[13,46,47] Similarly, for both the exchange part and the Coulombic part, i.e., the matrix elements in Eq. (8), a localized basis set would have reduced the scaling, as many exchange matrix elements would then vanish the grid extent, for the exchange integrals would be reduced. Once these improvements are made, the overall scaling of the method would reach a quadratic.[7,8]

Further work on this method will include fitting W to a given v_W interaction that goes beyond a translationally invariant interaction, but preserves the quasi-linear scaling of $\int v_W(r,r')\beta(r')dr'$. An improved fit would make it possible to use very few stochastic samplings of the difference operator $\{W - v_W\}$ or just forgo this term completely, keeping only v_W.

Further improvements include the anti-resonant to resonant transition couplings to go beyond the Tamm–Dancoff approximation. As mentioned in Sec. II D, since the contribution of this "off-diagonal" coupling in the BSE is substantially smaller than that of the resonant W, they could be represented by v_W alone rather than the full W, so, no additional W samplings would be needed.

In addition, dynamical corrections are needed in many systems with dominant $n \to \pi^*$ and $\pi \to \pi^*$ excitations.[48,49] While recent work has shown best results with a matrix perturbation theory method,[50] TDDFT-type approaches have been successful at capturing double excitations with a dynamical exchange kernel.[51–53]

While the formalism we present is naturally suitable to systems with fairly extended active (occupied and unoccupied) states, its extension to systems where the optically active states are highly localized, e.g., defects in otherwise homogeneous solids, should be straightforward. In such cases, as far as the action of W, the occupied (as well as the active unoccupied) state space should be divided into a few isolated states, while the rest of the states would be treated stochastically, as is done here. This will be developed in a future publication.

Finally, the grid-based approach presented here should be extendable to basis-set techniques, since in the BSE stage, it just uses the eigenstates and eigenvalues, and those come from an underlying DFT calculation, which could be grid-based (as is done here), or use basis-sets. Similarly, the action of stochastic W via a time-dependent Hartree propagation can also be done with basis-sets. The only fine points are that some of the intermediate calculations are most easily done on a grid, such as the calculation of the stochastic pair functions, $\beta(r)$, and the resulting convolutions; however, as these are mostly dependent on low wavelengths, rather rough grids could be employed, as is done in several basis-set techniques that use intermediate grids.

To summarize, our main point in this paper is that the optimized attenuated potential reduces the magnitude of the effective interaction W. This reduces the required number of stochastic samplings of $\{W - v_W\}$ to a manageable number, in the few thousands, enabling efficient BSE simulations. Furthermore, when fitting W to a translationally invariant interaction (which is applied by convolution), the resulting TDHF spectra, with v_W as the exchange interaction, are in quite good agreement with the exact W-based BSE results.

ACKNOWLEDGMENTS

We are grateful for the discussions with Vojtech Vleck. This work was 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 No. DE-SC0022198. N.C.B. acknowledges the National Science Foundation Graduate Research Fellowship Program, under Grant No. DGE-2034835. Computational resources were provided by the National Energy Research Scientific Computing Center, a DOE Office of Science user facility, supported by the Office of Science of the U.S. Department of Energy, under Contract No. DE-AC02-05CH11231, using NERSC Award No. BES-ERCAP0020089.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Nadine C. Bradbury: Conceptualization (lead); Data curation (lead); Formal analysis (lead); Investigation (lead); Methodology (equal); Software (equal); Validation (lead); Visualization (lead); Writing – original draft (lead). **Tucker Allen**: Data curation (equal); Validation (equal). **Minh Nguyen**: Data curation (equal). **Khaled Z. Ibrahim**: Methodology (equal); Software (equal). **Daniel Neuhauser**: Conceptualization (equal); Investigation (equal); Methodology (equal); Project administration (equal); Resources (equal); Software (equal); Supervision (equal); Validation (equal); Writing – original draft (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] X. Blase, I. Duchemin, D. Jacquemin, and P.-F. Loos, J. Phys. Chem. Lett. **11**, 7371 (2020).
[2] M. P. Ljungberg, P. Koval, F. Ferrari, D. Foerster, and D. Sánchez-Portal, Phys. Rev. B **92**, 075422 (2015).
[3] I. Duchemin and X. Blase, J. Chem. Phys. **150**, 174120 (2019).
[4] D. Rocca, Y. Ping, R. Gebauer, and G. Galli, Phys. Rev. B **85**, 045116 (2012).
[5] J. Deslippe, G. Samsonidze, D. A. Strubbe, M. Jain, M. L. Cohen, and S. G. Louie, Comput. Phys. Commun. **183**, 1269 (2012).
[6] D. Sangalli, A. Ferretti, H. Miranda, C. Attaccalite, I. Marri, E. Cannuccia, P. Melo, M. Marsili, F. Paleari, A. Marrazzo, G. Prandini, P. Bonfà, M. O. Atambo, F. Affinito, M. Palummo, A. Molina-Sánchez, C. Hogan, M. Grüning, D. Varsano, and A. Marini, J. Phys.: Condens. Matter **31**, 325902 (2019).
[7] A. Förster and L. Visscher, Front. Chem. **9**, 736591 (2021).
[8] A. Förster and L. Visscher, J. Chem. Theory Comput. **18**, 6779 (2022).
[9] N. C. Bradbury, M. Nguyen, J. R. Caram, and D. Neuhauser, J. Chem. Phys. **157**, 031104 (2022).
[10] F. Fuchs, C. Rödl, A. Schleife, and F. Bechstedt, Phys. Rev. B **78**, 085103 (2008).
[11] E. Rabani, R. Baer, and D. Neuhauser, Phys. Rev. B **91**, 235302 (2015).
[12] K. Begam, S. Bhandari, B. Maiti, and B. D. Dunietz, J. Chem. Theory Comput. **16**, 3287 (2020).
[13] D. Neuhauser, E. Rabani, Y. Cytter, and R. Baer, J. Phys. Chem. A **120**, 3071 (2015).
[14] Y. Gao, D. Neuhauser, R. Baer, and E. Rabani, J. Chem. Phys. **142**, 034106 (2015).
[15] J. W. Negele, Rev. Mod. Phys. **54**, 913 (1982).
[16] D. Neuhauser and R. Baer, J. Chem. Phys. **123**, 204105 (2005).
[17] A. Weiße, G. Wellein, A. Alvermann, and H. Fehske, Rev. Mod. Phys. **78**, 275 (2006).
[18] X. Gui, C. Holzer, and W. Klopper, J. Chem. Theory Comput. **14**, 2127 (2018).
[19] C. A. McKeon, S. M. Hamed, F. Bruneval, and J. B. Neaton, J. Chem. Phys. **157**, 074103 (2022).
[20] D. Neuhauser, Y. Gao, C. Arntsen, C. Karshenas, E. Rabani, and R. Baer, Phys. Rev. Lett. **113**, 076402 (2014).
[21] V. Vlček, E. Rabani, and D. Neuhauser, Phys. Rev. Mater. **2**, 030801(R) (2018).
[22] V. Vlček, R. Baer, E. Rabani, and D. Neuhauser, J. Chem. Phys. **149**, 174107 (2018).
[23] D. Neuhauser, J. Chem. Phys. **93**, 2611 (1990).
[24] V. A. Mandelshtam, T. R. Ravuri, and H. S. Taylor, Phys. Rev. Lett. **70**, 1932 (1993).
[25] R. Baer, D. Neuhauser, and E. Rabani, Annu. Rev. Phys. Chem. **73**, 255 (2022).
[26] D. Neuhauser, R. Baer, and E. Rabani, J. Chem. Phys. **141**, 041102 (2014).
[27] G. J. Martyna and M. E. Tuckerman, J. Chem. Phys. **110**, 2810 (1999).
[28] Y. Yang, E. R. Davidson, and W. Yang, Proc. Natl. Acad. Sci. U. S. A. **113**, E5098 (2016).
[29] M. B. Minameyer, Y. Xu, S. Frühwald, A. Görling, M. Delius, and T. Drewello, Chem. -Eur. J. **26**, 8729 (2020).
[30] E. Epifanovsky *et al.*, J. Chem. Phys. **155**, 084801 (2021).
[31] C. L. Reis, J. M. Pacheco, and J. L. Martins, Phys. Rev. B **68**, 155111 (2003).
[32] A. Willand, Y. O. Kvashnin, L. Genovese, Á. Vázquez-Mayagoitia, A. K. Deb, A. Sadeghi, T. Deutsch, and S. Goedecker, J. Chem. Phys. **138**, 104109 (2013).
[33] V. Vlček, W. Li, R. Baer, E. Rabani, and D. Neuhauser, Phys. Rev. B **98**, 075107 (2018).
[34] G. Onida, L. Reining, R. W. Godby, R. Del Sole, and W. Andreoni, Phys. Rev. Lett. **75**, 818 (1995).
[35] C. A. Rozzi, D. Varsano, A. Marini, E. K. U. Gross, and A. Rubio, Phys. Rev. B **73**, 205119 (2006).
[36] J. C. S. Costa, R. J. S. Taveira, C. F. R. A. C. Lima, A. Mendes, and L. M. N. B. F. Santos, Opt. Mater. **58**, 51 (2016).
[37] A. Menon, J. A. H. Dreyer, J. W. Martin, J. Akroyd, J. Robertson, and M. Kraft, Phys. Chem. Chem. Phys. **21**, 16240 (2019).
[38] C. Tönshoff and H. F. Bettinger, Angew. Chem., Int. Ed. **49**, 4125 (2010).
[39] J. Krüger, F. Eisenhut, J. M. Alonso, T. Lehmann, E. Guitián, D. Pérez, D. Skidin, F. Gamaleja, D. A. Ryndyk, C. Joachim, D. Peña, F. Moresco, and G. Cuniberti, Chem. Commun. **53**, 1583 (2017).
[40] C. Tönshoff and H. F. Bettinger, Chem. -Eur. J. **27**, 3193 (2020).
[41] R. Mondal, C. Tönshoff, D. Khon, D. C. Neckers, and H. F. Bettinger, J. Am. Chem. Soc. **131**, 14281 (2009).
[42] G. Rouillé, C. Jäger, M. Steglich, F. Huisken, T. Henning, G. Theumer, I. Bauer, and H.-J. Knölker, ChemPhysChem **9**, 2085 (2008).
[43] T. Rabenau, A. Simon, R. K. Kremer, and E. Sohmen, Z. Phys. B: Condens. Matter **90**, 69 (1993).
[44] R. W. Lof, M. A. van Veenendaal, H. T. Jonkman, and G. A. Sawatzky, J. Electron Spectrosc. Relat. Phenom. **72**, 83 (1995).
[45] T. Liu, C. Tonnelé, S. Zhao, L. Rondin, C. Elias, D. Medina-Lopez, H. Okuno, A. Narita, Y. Chassagneux, C. Voisin, S. Campidelli, D. Beljonne, and J.-S. Lauret, Nanoscale **14**, 3826 (2022).
[46] Y. Xu, R. Kaur, B. Wang, M. B. Minameyer, S. Gsänger, B. Meyer, T. Drewello, D. M. Guldi, and M. von Delius, J. Am. Chem. Soc. **140**, 13413 (2018).
[47] M. Romanova and V. Vlček, npj Comput. Mater. **8**, 11 (2022).
[48] Y. Ma, M. Rohlfing, and C. Molteni, Phys. Rev. B **80**, 241405 (2009).
[49] B. Baumeier, D. Andrienko, Y. Ma, and M. Rohlfing, J. Chem. Theory Comput. **8**, 997 (2012).
[50] P.-F. Loos and X. Blase, J. Chem. Phys. **153**, 114120 (2020).
[51] P. Romaniello, D. Sangalli, J. A. Berger, F. Sottile, L. G. Molinari, L. Reining, and G. Onida, J. Chem. Phys. **130**, 044108 (2009).
[52] M. Huix-Rotllant, A. Ipatov, A. Rubio, and M. E. Casida, Chem. Phys. **391**, 120 (2011).
[53] E. Rebolini and J. Toulouse, J. Chem. Phys. **144**, 094107 (2016).