FAST OPTICAL ABSORPTION SPECTRA CALCULATIONS FOR PERIODIC SOLID STATE SYSTEMS

Felix Henneke, Lin Lin, Christian Vorwerk, Claudia Draxl, Rupert Klein and Chao Yang

msp

msp

FAST OPTICAL ABSORPTION SPECTRA CALCULATIONS FOR PERIODIC SOLID STATE SYSTEMS

FELIX HENNEKE, LIN LIN, CHRISTIAN VORWERK,
CLAUDIA DRAXL, RUPERT KLEIN AND CHAO YANG

We present a method to construct an efficient approximation to the bare exchange and screened direct interaction kernels of the Bethe–Salpeter Hamiltonian for periodic solid state systems via the interpolative separable density fitting technique. We show that the cost of constructing the approximate Bethe–Salpeter Hamiltonian can be reduced to nearly optimal as $\mathbb{O}(N_k)$ with respect to the number of samples in the Brillouin zone N_k for the first time. In addition, we show that the cost for applying the Bethe–Salpeter Hamiltonian to a vector scales as $\mathbb{O}(N_k \log N_k)$. Therefore, the optical absorption spectrum, as well as selected excitation energies, can be efficiently computed via iterative methods such as the Lanczos method. This is a significant reduction from the $\mathbb{O}(N_k^2)$ and $\mathbb{O}(N_k^3)$ scaling associated with a brute force approach for constructing the Hamiltonian and diagonalizing the Hamiltonian, respectively. We demonstrate the efficiency and accuracy of this approach with both one-dimensional model problems and three-dimensional real materials (graphene and diamond). For the diamond system with $N_k = 2197$, it takes 6 hours to assemble the Bethe–Salpeter Hamiltonian and 4 hours to fully diagonalize the Hamiltonian using 169 cores when the brute force approach is used. The new method takes less than 3 minutes to set up the Hamiltonian and 24 minutes to compute the absorption spectrum on a single core.

1. Introduction

The Bethe–Salpeter equation (BSE), derived from the many-body perturbation theory (MBPT), is a widely used method for describing the optical absorption process in molecules and solids [32; 33; 36; 24; 1; 25; 7]. It models the behavior of an electron–hole pair, which is an excitation process with two quasiparticles. Solving the BSE requires constructing and diagonalizing a structured matrix, called the Bethe–Salpeter Hamiltonian (BSH). In the context of optical absorption, the

eigenvalues of the BSH are the exciton energies and the corresponding eigenfunctions yield the exciton wavefunctions. The BSH consists of the so-called bare exchange and screened direct interaction kernels that depend on single particle orbitals obtained from a quasiparticle (usually at the GW level) or mean-field calculation. For isolated systems such as molecules, the construction of these kernels requires at least $\mathbb{O}(N_e^5)$ operations in a conventional approach, where N_e is the number of electrons in the system. This is very costly for large systems that contain hundreds or more atoms. Recent efforts have actively explored methods for efficient representation of the BSH, in order to reduce the high computational cost of BSE calculations [4; 3; 16; 21; 30; 27; 28; 31; 23].

In a recent work [13], two of the authors have presented an efficient way to construct the BSH for molecular systems, and to efficiently solve the BSE eigenvalue problem using an iterative scheme. This approach is based on the recently developed interpolative separable density fitting (ISDF) decomposition [19; 20]. The ISDF decomposition has been applied to accelerate a number of applications in computational chemistry and materials science, including the computation of two-electron integrals [19], correlation energy in the random phase approximation [18], density functional perturbation theory [15], and hybrid density functional calculations [12; 8]. In this scheme, a matrix consisting of products of single particle orbital pairs is efficiently approximated as a low-rank matrix product of a matrix built with a small number of auxiliary basis vectors and an expansion coefficient matrix. This decomposition allows us to construct efficient representations of the bare exchange and screened direct kernels. For isolated molecular systems, the construction of the ISDF-compressed BSH matrix only requires $\mathbb{O}(N_e^3)$ operations when the rank of the numerical auxiliary basis is kept at $\mathbb{O}(N_e)$. This results in considerable reduction of the cost compared to the $\mathbb{O}(N_e^5)$ complexity required in a conventional approach. By keeping the interaction kernels in a decomposed form, the matrix–vector multiplications required in the iterative diagonalization procedures of the Hamiltonian H_{BSE} can be performed efficiently. We can further use these efficient matrix–vector multiplications in a structure-preserving Lanczos algorithm [34] to obtain an approximate absorption spectrum without an explicit diagonalization of the approximate H_{BSE}.

This paper generalizes the work in [13] to periodic solid state systems. According to the Bloch decomposition, each single particle orbital in a periodic system can be characterized by an orbital index i and a Brillouin zone index $\boldsymbol{k}$. Compared to isolated systems, the total number of electrons N_e is equal to the number of electrons per unit cell multiplied by the number of $\boldsymbol{k}$-points denoted by N_k. It has been observed that for many extended systems, the number of orbitals (both occupied and virtual orbitals) required for one particular $\boldsymbol{k}$ index can be relatively small, and is independent of N_e. Hence, the difficulty of optical absorption spectra

calculations for periodic systems mainly arise from the large number of k-points. This is particularly the case when the excitons are delocalized in the real space, or when the Fermi-surface is not smooth (such as graphene, and other metallic systems). In such case, N_k can often be rather large (from hundreds to hundreds of thousands; see, e.g., [29], where a $120 \times 120 \times 1$ k-grid is used for the quasi-two-dimensional MoS_2 system) in order to properly discretize and sample the Brillouin zone. The cost for constructing the bare exchange and screened direct kernels scales as $\mathbb{O}(N_k^2)$, while the cost for diagonalizing the corresponding BSH scales as $\mathbb{O}(N_k^3)$. This is prohibitively expensive when a dense discretization of the Brillouin zone is needed.

With the help of ISDF for periodic systems [20], we reduce the computational cost for producing optical absorption spectra to a scaling almost linear in N_k. First, the complexity of the bare exchange and screened direct kernel construction for extended systems is reduced to the optimal complexity of $\mathbb{O}(N_k)$. A sufficiently reduced representation of the pair product orbitals is possible, thanks to the smoothness of the single particle orbitals with respect to the k index, and the fact that the Brillouin zone is a compact domain. Second, the separable structure of the decomposition makes it possible to exploit a convolutional structure in the screened direct kernel. The complexity of applying the approximated kernels to a vector with respect to N_k is thus only $\mathbb{O}(N_k \log N_k)$. Instead of diagonalizing the BSH directly, we use iterative methods such as the Lanczos method to evaluate the optical absorption spectrum. The same strategy can be applied to evaluate selected excitation energies.

Despite the increasingly wide adoption of the BSE theory in condensed matter physics and quantum chemistry for analyzing optical properties of materials, we could not find a precise mathematical description of how the BSH is constructed for periodic systems in the literature. Therefore, after concise review of the single particle theory and the Bethe–Salpeter equation for periodic systems in Section 2.1, we provide a relatively self-contained derivation of the BSE for periodic systems in Section 2.2 from a numerical linear algebra perspective. We hope our presentation (especially using a discretized Brillouin zone so that all matrices are of finite dimension) is useful to readers not familiar with the matter.

Then the rest of the paper is organized as follows. The interpolative separable density fitting for periodic systems is introduced in Section 3, and the application of the approximate BSH in the ISDF format to a vector in Section 4. The numerical results are presented in Section 5, followed by a conclusion in Section 6.

2. Preliminaries

2.1. *Single particle theory for periodic systems.* To facilitate further discussion we briefly review Bloch–Floquet theory for periodic systems. Without loss of generality we consider a three-dimensional crystal. The *Bravais lattice* with lattice

vectors $a_1, a_2, a_3 \in \mathbb{R}^3$ is defined as

$$\mathbb{L} = \{R \mid R = n_1 a_1 + n_2 a_2 + n_3 a_3, \ n_1, n_2, n_3 \in \mathbb{Z}\}. \tag{2-1}$$

In single particle theories such as the Kohn–Sham density functional theory, the self-consistent effective potential V_{eff} is real-valued and $\mathbb{L}$-periodic, i.e.,

$$V_{\text{eff}}(r + R) = V_{\text{eff}}(r) \quad \text{for all } r \in \mathbb{R}^3 \text{ and } R \in \mathbb{L}.$$

The unit cell is defined as

$$\Omega = \{r = c_1 a_1 + c_2 a_2 + c_3 a_3 \mid 0 \le c_1, c_2, c_3 < 1\}. \tag{2-2}$$

The Bravais lattice induces a reciprocal lattice $\mathbb{L}^*$, with its lattice vectors b_1, b_2, b_3 satisfying $a_\alpha \cdot b_\beta = 2\pi \delta_{\alpha\beta}$, $\alpha, \beta \in \{1, 2, 3\}$. The unit cell of the reciprocal lattice is called the (first) Brillouin zone and denoted by Ω^*, defined as

$$\Omega^* = \left\{ k = k_1 b_1 + k_2 b_2 + k_3 b_3 \mid -\tfrac{1}{2} \le k_1, k_2, k_3 < \tfrac{1}{2} \right\}.$$

The Brillouin zone has a number of special points related to the symmetry of the crystal. The common special point is the Γ-point, which corresponds to $k = [0, 0, 0]^\top$.

According to the Bloch–Floquet theory, the spectrum of the Hamiltonian $\mathcal{H} = -\tfrac{1}{2}\nabla_r^2 + V_{\text{eff}}(r)$ can be relabeled using two indices (i, k), where $i \in \mathbb{N}$ is called the band index and $k \in \Omega^*$ is the Brillouin zone index. Each generalized eigenfunction $\psi_{ik}(r)$ is known as a Bloch orbital and satisfies $\mathcal{H}\psi_{ik}(r) = \epsilon_{ik}\psi_{ik}(r)$ with Bloch boundary conditions $\psi_{ik}(r + R) = e^{ik \cdot R}\psi_{ik}(r)$ for any $R \in \mathbb{L}$. Furthermore, ψ_{ik} can be decomposed using the Bloch decomposition

$$\psi_{ik}(r) = e^{ik \cdot r} u_{ik}(r), \tag{2-3}$$

where $u_{ik}(r)$ is the periodic part of $\psi_{ik}(r)$ satisfying the periodic boundary condition on the unit cell

$$u_{ik}(r + R) = u_{ik}(r) \quad \text{for all } R \in \mathbb{L}. \tag{2-4}$$

It can be directly obtained by solving the eigenvalue problem

$$\mathcal{H}(k)u_{ik} = \epsilon_{ik}u_{ik}(r), \quad r \in \Omega, \ k \in \Omega^*, \tag{2-5}$$

where $\mathcal{H}(k) = -\tfrac{1}{2}(\nabla_r + ik)^2 + V_{\text{eff}}(r)$. For each $k \in \Omega^*$, the eigenvalues ϵ_{ik} are ordered nondecreasingly. For a fixed i, $\{\epsilon_{ik}\}$ as a function of k is called a *Bloch band*. The collection of all eigenvalues forms the *band structure* of the crystal, which characterizes the spectrum of the operator $\mathcal{H}$.

In the discussion below, we denote by N_v the number of valence bands (i.e., occupied orbitals per unit cell in the ground state) and N_c the number of conduction bands (i.e., unoccupied orbitals per unit cell in the ground state). We also define

$N = N_v + N_c$. We assume the systems to be insulating, in the sense that the following band isolation conditions between the valence and conduction bands are satisfied:

$$\inf|\epsilon_{ik} - \epsilon_{i'k'}| := \epsilon_g > 0, \quad k, k' \in \Omega^*, \ 1 \le i \le N_v, \ N_v + 1 \le i' \le N. \quad (2\text{-}6)$$

Denote by $|\Omega|$ the volume of the unit cell, and by

$$|\Omega^*| = \frac{(2\pi)^3}{|\Omega|}$$

the volume of the Brillouin zone. The Bloch orbitals $\{\psi_{ik}\}$ satisfy the orthonormality condition in the distributional sense:

$$\int_{\mathbb{R}^3} \psi_{i'k'}^*(r)\psi_{i,k}(r)\,\mathrm{d}r = |\Omega^*|\delta_{i',i}\delta(k' - k). \quad (2\text{-}7)$$

Here $\delta_{i',i}$ is the Kronecker δ symbol for a discrete set, while $\delta(k' - k)$ is the Dirac delta distribution. Equation (2-7) implies the normalization condition when integrated over the Brillouin zone:

$$\frac{1}{|\Omega^*|} \int_{\Omega^*} \int_{\mathbb{R}^3} \psi_{i'k}^*(r)\psi_{ik}(r)\,\mathrm{d}r\,\mathrm{d}k = \delta_{i',i}. \quad (2\text{-}8)$$

From the Bloch orbitals, the ground state electron density can be constructed as

$$\rho(r) = \frac{1}{|\Omega^*|} \int_{\Omega^*} \sum_{i=1}^{N_v} |\psi_{ik}(r)|^2 \,\mathrm{d}k = \frac{1}{|\Omega^*|} \int_{\Omega^*} \sum_{i=1}^{N_v} |u_{ik}(r)|^2 \,\mathrm{d}k. \quad (2\text{-}9)$$

In order to practically perform calculations for periodic systems, the integration with respect to the Brillouin zone Ω^* needs to be discretized using a quadrature. The most commonly used scheme is based on the Monkhorst–Pack grid [22]

$$\mathcal{K}_s^\ell = \left\{ \sum_{\alpha=1}^{3} \frac{m_\alpha - s_\alpha}{N_\alpha^\ell} b_\alpha \ \middle| \ m_\alpha = -\frac{N_\alpha^\ell}{2}+1, \ldots, \frac{N_\alpha^\ell}{2}, \ 0 \le s_\alpha < 1, \ \alpha = 1, 2, 3 \right\}. \quad (2\text{-}10)$$

It is clear that $\mathcal{K}_s^\ell \subset \Omega^*$ and that it corresponds to a uniform discretization of the Brillouin zone. When the shift vector $s = 0$, we denote $\mathcal{K}^\ell := \mathcal{K}_0^\ell$, and the calculation of periodic systems can be *equivalently* performed using a supercell consisting of $N_1^\ell \times N_2^\ell \times N_3^\ell$ unit cells. The supercell is denoted by Ω^ℓ, and is further equipped with a periodic boundary condition called the Born–von Karman boundary condition [2]. The calculation of a periodic crystal can thus be recovered by taking the limit $N_\alpha^\ell \to \infty$. We denote by $N_k \equiv N^\ell := N_1^\ell N_2^\ell N_3^\ell$ the total number of unit cells, or equivalently the total number of Monkhorst–Pack grid points in the Brillouin zone.

Assuming the Brillouin zone is discretized using $\mathcal{K}^\ell$, the orthogonality condition (2-7) becomes

$$\int_{\Omega^\ell} \psi_{i'\boldsymbol{k}'}^*(\boldsymbol{r})\psi_{i\boldsymbol{k}}(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r} = \delta_{i',i}\delta_{\boldsymbol{k}',\boldsymbol{k}}, \quad \boldsymbol{k},\boldsymbol{k}' \in \mathcal{K}^\ell. \tag{2-11}$$

We also modify the Bloch decomposition as

$$\psi_{i\boldsymbol{k}}(\boldsymbol{r}) = \frac{1}{\sqrt{N^\ell}}e^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}u_{i\boldsymbol{k}}(\boldsymbol{r}), \quad \boldsymbol{k} \in \mathcal{K}^\ell. \tag{2-12}$$

Here the normalization factor $1/\sqrt{N^\ell}$ is introduced so that the orthogonality condition for the periodic part implies

$$\int_\Omega u_{i'\boldsymbol{k}}^*(\boldsymbol{r})u_{i\boldsymbol{k}}(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r} = \delta_{i',i}, \quad \boldsymbol{k} \in \mathcal{K}^\ell. \tag{2-13}$$

To facilitate the bookkeeping effort of various relevant constants in practical calculations, in the discussion below we will always assume that the Brillouin zone is discretized into $\mathcal{K}^\ell$ with a corresponding supercell Ω^ℓ. The volume of the supercell is $|\Omega^\ell| = N^\ell|\Omega| = N_k|\Omega|$. The unit cell is further discretized into a uniform grid $\{\boldsymbol{r}_i\}_{i=1}^{N_g}$. Practical BSE calculations often truncate the number of conduction bands aggressively, in the sense that $N_g \gg N_v + N_c =: N$. Numerical results indicate that in many cases, the low-lying excitation spectrum is relatively insensitive to N_c, and one can often choose $N_c \approx N_v$. Unless otherwise clarified, we may not distinguish a continuous vector $u(\boldsymbol{r})$ and the corresponding discretized vector $\{u(\boldsymbol{r}_i)\}$. Similarly, when the context is clear, we do not distinguish the kernel of an operator $A(\boldsymbol{r},\boldsymbol{r}')$ and its discretized matrix $\{A(\boldsymbol{r}_i,\boldsymbol{r}_j)\}$.

2.2. *Bethe–Salpeter equation for periodic systems.* The Bethe–Salpeter equation is an eigenvalue problem of the form

$$H_{\mathrm{BSE}}X = EX, \tag{2-14}$$

where H_{BSE} is the Bethe–Salpeter Hamiltonian (BSH), X is the exciton wavefunction, and E is the corresponding exciton energy. For periodic systems, the BSH has the block structure

$$H_{\mathrm{BSE}} = \begin{bmatrix} D + 2V_A - W_A & 2V_B - W_B \\ -2\overline{V}_B + \overline{W}_B & -D - 2\overline{V}_A + \overline{W}_A \end{bmatrix}, \tag{2-15}$$

where $D(i_v i_c \boldsymbol{k}, j_v j_c \boldsymbol{k}') = (\epsilon_{i_c\boldsymbol{k}} - \epsilon_{i_v\boldsymbol{k}})\delta_{i_v,j_v}\delta_{i_c,j_c}\delta_{\boldsymbol{k},\boldsymbol{k}'}$ is an $(N_v N_c N_k) \times (N_v N_c N_k)$ diagonal matrix. The quasiparticle energies $\epsilon_{i_v\boldsymbol{k}}, \epsilon_{i_c\boldsymbol{k}}$ are typically obtained from a GW calculation [32]. The V_A and V_B matrices represent the bare *exchange* interaction of electron–hole pairs, and the W_A and W_B matrices are referred to as

the screened *direct* interaction of electron–hole pairs. These matrices are defined as

$$V_A(i_v i_c \boldsymbol{k}, j_v j_c \boldsymbol{k}') = \int_{\Omega^\ell \times \Omega^\ell} \overline{\psi}_{i_c \boldsymbol{k}}(\boldsymbol{r}) \psi_{i_v \boldsymbol{k}}(\boldsymbol{r}) V(\boldsymbol{r}, \boldsymbol{r}') \overline{\psi}_{j_v \boldsymbol{k}'}(\boldsymbol{r}') \psi_{j_c \boldsymbol{k}'}(\boldsymbol{r}') \, \mathrm{d}\boldsymbol{r} \, \mathrm{d}\boldsymbol{r}',$$

$$V_B(i_v i_c \boldsymbol{k}, j_v j_c \boldsymbol{k}') = \int_{\Omega^\ell \times \Omega^\ell} \overline{\psi}_{i_c \boldsymbol{k}}(\boldsymbol{r}) \psi_{i_v \boldsymbol{k}}(\boldsymbol{r}) V(\boldsymbol{r}, \boldsymbol{r}') \overline{\psi}_{j_c \boldsymbol{k}'}(\boldsymbol{r}') \psi_{j_v \boldsymbol{k}'}(\boldsymbol{r}') \, \mathrm{d}\boldsymbol{r} \, \mathrm{d}\boldsymbol{r}',$$

$$W_A(i_v i_c \boldsymbol{k}, j_v j_c \boldsymbol{k}') = \int_{\Omega^\ell \times \Omega^\ell} \overline{\psi}_{i_c \boldsymbol{k}}(\boldsymbol{r}) \psi_{j_c \boldsymbol{k}'}(\boldsymbol{r}) W(\boldsymbol{r}, \boldsymbol{r}') \overline{\psi}_{j_v \boldsymbol{k}'}(\boldsymbol{r}') \psi_{i_v \boldsymbol{k}}(\boldsymbol{r}') \, \mathrm{d}\boldsymbol{r} \, \mathrm{d}\boldsymbol{r}', \tag{2-16}$$

$$W_B(i_v i_c \boldsymbol{k}, j_v j_c \boldsymbol{k}') = \int_{\Omega^\ell \times \Omega^\ell} \overline{\psi}_{i_c \boldsymbol{k}}(\boldsymbol{r}) \psi_{j_v \boldsymbol{k}'}(\boldsymbol{r}) W(\boldsymbol{r}, \boldsymbol{r}') \overline{\psi}_{j_c \boldsymbol{k}'}(\boldsymbol{r}') \psi_{i_v \boldsymbol{k}}(\boldsymbol{r}') \, \mathrm{d}\boldsymbol{r} \, \mathrm{d}\boldsymbol{r}'.$$

Here $\psi_{i_v \boldsymbol{k}}$ and $\psi_{i_c \boldsymbol{k}}$ are the valence and conduction single particle orbitals typically obtained from a Kohn–Sham density functional theory (KSDFT) calculation, respectively, and $V(\boldsymbol{r}, \boldsymbol{r}')$ and $W(\boldsymbol{r}, \boldsymbol{r}')$ are the bare and screened Coulomb interactions. Both V_A and W_A are Hermitian, whereas V_B and W_B are complex symmetric. Within the so-called Tamm–Dancoff approximation (TDA) [25], both V_B and W_B are neglected in (2-15). In this case, the H_{BSE} becomes Hermitian and we can focus on computing the upper left block of H_{BSE}. Both the KSDFT and GW calculations can be challenging in their own right. In this work, however, we consider their output as given and the starting point of our BSE calculation.

In the following discussion, when a single index i is used, it refers to either i_v or i_c. Using the Bloch decomposition (2-12), the matrix elements of the BSH can be written using the periodic part of the orbitals as

$$V_A(i_v i_c \boldsymbol{k}, j_v j_c \boldsymbol{k}') = \frac{1}{N_k^2} \int_{\Omega^\ell \times \Omega^\ell} \overline{u}_{i_c \boldsymbol{k}}(\boldsymbol{r}) u_{i_v \boldsymbol{k}}(\boldsymbol{r}) V(\boldsymbol{r}, \boldsymbol{r}') \overline{u}_{j_v \boldsymbol{k}'}(\boldsymbol{r}') u_{j_c \boldsymbol{k}'}(\boldsymbol{r}') \, \mathrm{d}\boldsymbol{r} \, \mathrm{d}\boldsymbol{r}',$$

$$V_B(i_v i_c \boldsymbol{k}, j_v j_c \boldsymbol{k}') = \frac{1}{N_k^2} \int_{\Omega^\ell \times \Omega^\ell} \overline{u}_{i_c \boldsymbol{k}}(\boldsymbol{r}) u_{i_v \boldsymbol{k}}(\boldsymbol{r}) V(\boldsymbol{r}, \boldsymbol{r}') \overline{u}_{j_c \boldsymbol{k}'}(\boldsymbol{r}') u_{j_v \boldsymbol{k}'}(\boldsymbol{r}') \, \mathrm{d}\boldsymbol{r} \, \mathrm{d}\boldsymbol{r}',$$

$$W_A(i_v i_c \boldsymbol{k}, j_v j_c \boldsymbol{k}') = \frac{1}{N_k^2} \int_{\Omega^\ell \times \Omega^\ell} e^{-\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}') \cdot (\boldsymbol{r}-\boldsymbol{r}')} \overline{u}_{i_c \boldsymbol{k}}(\boldsymbol{r}) u_{j_c \boldsymbol{k}'}(\boldsymbol{r}) \tag{2-17}$$
$$\times \, W(\boldsymbol{r}, \boldsymbol{r}') \overline{u}_{j_v \boldsymbol{k}'}(\boldsymbol{r}') u_{i_v \boldsymbol{k}}(\boldsymbol{r}') \, \mathrm{d}\boldsymbol{r} \, \mathrm{d}\boldsymbol{r}',$$

$$W_B(i_v i_c \boldsymbol{k}, j_v j_c \boldsymbol{k}') = \frac{1}{N_k^2} \int_{\Omega^\ell \times \Omega^\ell} e^{-\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}') \cdot (\boldsymbol{r}-\boldsymbol{r}')} \overline{u}_{i_c \boldsymbol{k}}(\boldsymbol{r}) u_{j_v \boldsymbol{k}'}(\boldsymbol{r})$$
$$\times \, W(\boldsymbol{r}, \boldsymbol{r}') \overline{u}_{j_c \boldsymbol{k}'}(\boldsymbol{r}') u_{i_v \boldsymbol{k}}(\boldsymbol{r}') \, \mathrm{d}\boldsymbol{r} \, \mathrm{d}\boldsymbol{r}'.$$

Note that V_A, V_B in (2-17) do not involve the phase factors, since the factor $e^{\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{r}}$ cancels exactly due to the complex conjugate operation. The phase factor only appears in the W_A, W_B terms.

Equation (2-17) requires the evaluation of integrals of the form

$$\mathcal{V}(f, g) := \frac{1}{N_k} \int_{\Omega^\ell \times \Omega^\ell} \overline{f}(\boldsymbol{r}) V(\boldsymbol{r}, \boldsymbol{r}') g(\boldsymbol{r}') \, \mathrm{d}\boldsymbol{r} \, \mathrm{d}\boldsymbol{r}', \tag{2-18}$$

and

$$\mathcal{W}_{\boldsymbol{q}}(f, g) := \frac{1}{N_k} \int_{\Omega^\ell \times \Omega^\ell} e^{-\mathrm{i}\boldsymbol{q}\cdot(\boldsymbol{r}-\boldsymbol{r}')} \bar{f}(\boldsymbol{r}) W(\boldsymbol{r}, \boldsymbol{r}') g(\boldsymbol{r}') \, \mathrm{d}\boldsymbol{r} \, \mathrm{d}\boldsymbol{r}'. \qquad (2\text{-}19)$$

Using such notation,

$$
\begin{aligned}
V_A(i_v i_c \boldsymbol{k}, j_v j_c \boldsymbol{k}') &= \frac{1}{N_k} \mathcal{V}(\bar{u}_{i_v \boldsymbol{k}} u_{i_c \boldsymbol{k}}, \bar{u}_{j_v \boldsymbol{k}'} u_{j_c \boldsymbol{k}'}), \\
V_B(i_v i_c \boldsymbol{k}, j_v j_c \boldsymbol{k}') &= \frac{1}{N_k} \mathcal{V}(\bar{u}_{i_v \boldsymbol{k}} u_{i_c \boldsymbol{k}}, \bar{u}_{j_c \boldsymbol{k}'} u_{j_v \boldsymbol{k}'}), \\
W_A(i_v i_c \boldsymbol{k}, j_v j_c \boldsymbol{k}') &= \frac{1}{N_k} \mathcal{W}_{\boldsymbol{k}-\boldsymbol{k}'}(\bar{u}_{j_c \boldsymbol{k}'} u_{i_c \boldsymbol{k}}, \bar{u}_{j_v \boldsymbol{k}'} u_{i_v \boldsymbol{k}}), \\
W_B(i_v i_c \boldsymbol{k}, j_v j_c \boldsymbol{k}') &= \frac{1}{N_k} \mathcal{W}_{\boldsymbol{k}-\boldsymbol{k}'}(\bar{u}_{j_v \boldsymbol{k}'} u_{i_c \boldsymbol{k}}, \bar{u}_{j_c \boldsymbol{k}'} u_{i_v \boldsymbol{k}}).
\end{aligned}
\qquad (2\text{-}20)
$$

In (2-18) and (2-19), f, g are periodic functions in the unit cell, and can be represented using their Fourier representations. For instance,

$$f(\boldsymbol{r}) = \sum_{\boldsymbol{G} \in \mathbb{L}^*} \hat{f}(\boldsymbol{G}) e^{\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}}, \qquad (2\text{-}21)$$

and its Fourier coefficients can be computed as

$$\hat{f}(\boldsymbol{G}) = \frac{1}{|\Omega|} \int_\Omega e^{-\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}} f(\boldsymbol{r}) \, \mathrm{d}\boldsymbol{r}. \qquad (2\text{-}22)$$

Hence, Parseval's identity reads

$$\int_\Omega \bar{f}(\boldsymbol{r}) g(\boldsymbol{r}) \, \mathrm{d}\boldsymbol{r} = |\Omega| \sum_{\boldsymbol{G} \in \mathbb{L}^*} \bar{\hat{f}}(\boldsymbol{G}) \hat{g}(\boldsymbol{G}). \qquad (2\text{-}23)$$

Both of the kernels V, W satisfy the translation symmetry

$$V(\boldsymbol{r}+\boldsymbol{R}, \boldsymbol{r}'+\boldsymbol{R}) = V(\boldsymbol{r}, \boldsymbol{r}'), \quad W(\boldsymbol{r}+\boldsymbol{R}, \boldsymbol{r}'+\boldsymbol{R}) = W(\boldsymbol{r}, \boldsymbol{r}') \text{ for all } \boldsymbol{R} \in \mathbb{L}. \quad (2\text{-}24)$$

Equation (2-24) also defines the values of V, W for $\boldsymbol{r}, \boldsymbol{r}'$ beyond the supercell Ω^ℓ. The Fourier representation of V takes the form

$$V(\boldsymbol{r}, \boldsymbol{r}') = \frac{1}{|\Omega^\ell|} \sum_{\boldsymbol{k} \in \mathcal{K}^\ell} \sum_{\boldsymbol{G}, \boldsymbol{G}'} e^{\mathrm{i}(\boldsymbol{k}+\boldsymbol{G})\cdot\boldsymbol{r}} \widehat{V}_{\boldsymbol{k}}(\boldsymbol{G}, \boldsymbol{G}') e^{-\mathrm{i}(\boldsymbol{k}+\boldsymbol{G}')\cdot\boldsymbol{r}'}, \qquad (2\text{-}25)$$

and the Fourier coefficients can be computed as

$$\widehat{V}_{\boldsymbol{k}}(\boldsymbol{G}, \boldsymbol{G}') = \frac{1}{|\Omega^\ell|} \int_{\Omega^\ell \times \Omega^\ell} \mathrm{d}\boldsymbol{r} \, \mathrm{d}\boldsymbol{r}' \, e^{-\mathrm{i}(\boldsymbol{k}+\boldsymbol{G})\cdot\boldsymbol{r}} V(\boldsymbol{r}, \boldsymbol{r}') e^{\mathrm{i}(\boldsymbol{k}+\boldsymbol{G}')\cdot\boldsymbol{r}'}. \qquad (2\text{-}26)$$

Similarly, the Fourier representation for W can be defined.

It should be noted that the Coulomb kernel V only depends on the distance between r and r', i.e., it has the further translational symmetry property that

$$V(r + r'', r' + r'') = V(r, r') \quad \text{for all } r'' \in \Omega^\ell. \tag{2-27}$$

As a result, its Fourier transform $\widehat{V}_k(G, G')$ can be simplified into a diagonal matrix

$$\widehat{V}_k(G, G') = \frac{4\pi}{|k + G|^2} \delta_{G, G'}. \tag{2-28}$$

In fact, the Coulomb kernel periodized with respect to the supercell Ω^ℓ is defined to be the inverse Fourier transform of (2-28).

Using such notation, we have

$$\int_{\Omega^\ell} V(r, r') g(r') \, dr'$$

$$= \frac{1}{|\Omega^\ell|} \int_{\Omega^\ell} dr' \sum_{k \in \mathcal{K}^\ell} \sum_{G, G'} e^{i(k+G)\cdot r} \widehat{V}_k(G, G') e^{-i(k+G')\cdot r'} g(r')$$

$$= \frac{1}{|\Omega^\ell|} \sum_{R \in \mathbb{L}} \int_{\Omega} dr' \sum_{k \in \mathcal{K}^\ell} \sum_{G, G'} e^{i(k+G)\cdot r} \widehat{V}_k(G, G') e^{-i(k+G')\cdot (r'+R)} g(r' + R)$$

$$= \frac{1}{|\Omega^\ell|} \int_{\Omega} dr' \sum_{k \in \mathcal{K}^\ell} \sum_{R \in \mathbb{L}} e^{-ik\cdot R} \sum_{G, G'} e^{i(k+G)\cdot r} \widehat{V}_k(G, G') e^{-i(k+G')\cdot r'} g(r'). \tag{2-29}$$

Here we have used $e^{-iG'\cdot R} = 1$ and the fact that g is periodic with respect to the unit cell Ω, as well as the identity

$$\int_{\Omega^\ell} f(r') \, dr' = \sum_{R \in \mathbb{L}} \int_{\Omega} f(r' + R) \, dr'. \tag{2-30}$$

Furthermore, from (2-22) and the identity

$$\sum_{R \in \mathbb{L}} e^{-ik\cdot R} = N_k \delta_{k, 0}$$

we have

$$\int_{\Omega^\ell} V(r, r') g(r') \, dr' = \frac{1}{|\Omega|} \int_{\Omega} dr' \sum_{G, G'} e^{iG\cdot r} \widehat{V}_0(G, G') e^{-iG'\cdot r'} g(r')$$

$$= \sum_{G, G'} e^{iG\cdot r} \widehat{V}_0(G, G') \hat{g}(G'). \tag{2-31}$$

Compared to (2-28), the definition of $\widehat{V}_0$ should be modified to

$$\widehat{V}_0(G, G') = \begin{cases} (4\pi/|G|^2) \delta_{G, G'}, & G \neq 0, \\ 0, & G = 0. \end{cases} \tag{2-32}$$

Another way to understand (2-32) is that it can only be applied to a mean-zero function $g(\boldsymbol{r})$, such that $\hat{g}(\boldsymbol{0}) = 0$. In other words, g should be in the range of the Laplacian operator with the periodic boundary condition. This is indeed correct for BSE calculations, due to the orthogonality condition between the valence and conduction bands

$$\int_{\Omega} \bar{u}_{i_c \boldsymbol{k}}(\boldsymbol{r}) u_{i_v \boldsymbol{k}}(\boldsymbol{r}) \, \mathrm{d}\boldsymbol{r} = 0.$$

This implies

$$
\begin{aligned}
\mathcal{V}(f, g) &= \frac{1}{N_k} \int_{\Omega^\ell} \bar{f}(\boldsymbol{r}) \sum_{\boldsymbol{G}, \boldsymbol{G}'} e^{\mathrm{i}\boldsymbol{G} \cdot \boldsymbol{r}} \widehat{V}_0(\boldsymbol{G}, \boldsymbol{G}') \hat{g}(\boldsymbol{G}') \\
&= \int_{\Omega} \bar{f}(\boldsymbol{r}) \sum_{\boldsymbol{G}, \boldsymbol{G}'} e^{\mathrm{i}\boldsymbol{G} \cdot \boldsymbol{r}} \widehat{V}_0(\boldsymbol{G}, \boldsymbol{G}') \hat{g}(\boldsymbol{G}') \\
&= |\Omega| \sum_{\boldsymbol{G}, \boldsymbol{G}'} \bar{\hat{f}}(\boldsymbol{G}) \widehat{V}_0(\boldsymbol{G}, \boldsymbol{G}') \hat{g}(\boldsymbol{G}') \\
&= |\Omega| \sum_{\boldsymbol{G} \neq 0} \frac{4\pi}{|\boldsymbol{G}|^2} \bar{\hat{f}}(\boldsymbol{G}) \hat{g}(\boldsymbol{G}).
\end{aligned}
\tag{2-33}
$$

Similarly for the W part,

$$
\begin{aligned}
&\int_{\Omega^\ell} e^{-\mathrm{i}\boldsymbol{q} \cdot (\boldsymbol{r} - \boldsymbol{r}')} W(\boldsymbol{r}, \boldsymbol{r}') g(\boldsymbol{r}') \, \mathrm{d}\boldsymbol{r}' \\
&= \frac{1}{|\Omega^\ell|} \int_{\Omega^\ell} \mathrm{d}\boldsymbol{r}' e^{-\mathrm{i}\boldsymbol{q} \cdot (\boldsymbol{r} - \boldsymbol{r}')} \sum_{\boldsymbol{k} \in \mathcal{K}^\ell} \sum_{\boldsymbol{G}, \boldsymbol{G}'} e^{\mathrm{i}(\boldsymbol{k} + \boldsymbol{G}) \cdot \boldsymbol{r}} \widehat{W}_{\boldsymbol{k}}(\boldsymbol{G}, \boldsymbol{G}') e^{-\mathrm{i}(\boldsymbol{k} + \boldsymbol{G}') \cdot \boldsymbol{r}'} g(\boldsymbol{r}') \\
&= \frac{1}{|\Omega^\ell|} \int_{\Omega} \mathrm{d}\boldsymbol{r}' e^{\mathrm{i}(\boldsymbol{k} - \boldsymbol{q}) \cdot (\boldsymbol{r} - \boldsymbol{r}')} \sum_{\boldsymbol{k} \in \mathcal{K}^\ell} \sum_{\boldsymbol{R} \in \mathbb{L}} e^{-\mathrm{i}(\boldsymbol{k} - \boldsymbol{q}) \cdot \boldsymbol{R}} \\
&\qquad\qquad\qquad\qquad\qquad \times \sum_{\boldsymbol{G}, \boldsymbol{G}'} e^{\mathrm{i}\boldsymbol{G} \cdot \boldsymbol{r}} \widehat{W}_{\boldsymbol{k}}(\boldsymbol{G}, \boldsymbol{G}') e^{-\mathrm{i}\boldsymbol{G}' \cdot \boldsymbol{r}'} g(\boldsymbol{r}').
\end{aligned}
\tag{2-34}
$$

In order to obtain a nonvanishing quantity in the equation above, note that the quantity $\sum_{\boldsymbol{R} \in \mathbb{L}} e^{-\mathrm{i}(\boldsymbol{k} - \boldsymbol{q}) \cdot \boldsymbol{R}} = N_k$ if $\boldsymbol{k} - \boldsymbol{q} \in \mathbb{L}^*$, and is otherwise 0. Therefore, the summation with respect to $\boldsymbol{k}$ should be restricted to those satisfying

$$\boldsymbol{k} - \boldsymbol{q} = \boldsymbol{G}'', \quad \boldsymbol{G}'' \in \mathbb{L}^*.$$

Since $\boldsymbol{k}$ is restricted to the first Brillouin zone, there is a unique $\boldsymbol{G}''$ (and therefore $\boldsymbol{k}$) for each given $\boldsymbol{q}$ satisfying this relation. Also note that $\boldsymbol{k} - \boldsymbol{q}$ may exceed the first Brillouin zone. In other words, it is indeed possible to have $\boldsymbol{G}'' \neq \boldsymbol{0}$. Then for a

given $\boldsymbol{q}$,

$$
\int_{\Omega^\ell} e^{-\mathrm{i}\boldsymbol{q}\cdot(\boldsymbol{r}-\boldsymbol{r}')} W(\boldsymbol{r},\boldsymbol{r}') g(\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}'
$$

$$
= \frac{1}{|\Omega|}\int_\Omega \mathrm{d}\boldsymbol{r}' \sum_{\boldsymbol{G},\boldsymbol{G}'} e^{\mathrm{i}(\boldsymbol{G}+\boldsymbol{G}'')\cdot\boldsymbol{r}} \widehat{W}_{\boldsymbol{G}''+\boldsymbol{q}}(\boldsymbol{G},\boldsymbol{G}') e^{-\mathrm{i}(\boldsymbol{G}'+\boldsymbol{G}'')\cdot\boldsymbol{r}'} g(\boldsymbol{r}')
$$

$$
= \sum_{\boldsymbol{G},\boldsymbol{G}'} e^{\mathrm{i}(\boldsymbol{G}+\boldsymbol{G}'')\cdot\boldsymbol{r}} \widehat{W}_{\boldsymbol{G}''+\boldsymbol{q}}(\boldsymbol{G},\boldsymbol{G}') \hat{g}(\boldsymbol{G}'+\boldsymbol{G}'')
$$

$$
= \sum_{\boldsymbol{G},\boldsymbol{G}'} e^{\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}} \widehat{W}_{\boldsymbol{G}''+\boldsymbol{q}}(\boldsymbol{G}-\boldsymbol{G}'',\boldsymbol{G}'-\boldsymbol{G}'') \hat{g}(\boldsymbol{G}')
$$

$$
= \sum_{\boldsymbol{G},\boldsymbol{G}'} e^{\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}} \widehat{W}_{\boldsymbol{q}}(\boldsymbol{G},\boldsymbol{G}') \hat{g}(\boldsymbol{G}'). \tag{2-35}
$$

In the last equality, we have used the definition of the Fourier coefficients in (2-26). We then readily have

$$
\mathcal{W}_{\boldsymbol{q}}(f,g) = |\Omega| \sum_{\boldsymbol{G},\boldsymbol{G}'} \bar{\hat{f}}(\boldsymbol{G}) \widehat{W}_{\boldsymbol{q}}(\boldsymbol{G},\boldsymbol{G}') \hat{g}(\boldsymbol{G}'). \tag{2-36}
$$

Therefore, despite that $\mathcal{W}_{\boldsymbol{q}}(f,g)$ is significantly more complex to define, the resulting formula in the Fourier representation is remarkably similar to the form of $\mathcal{V}(f,g)$.

3. Interpolative separable density fitting for periodic systems

In order to reduce the computational complexity, we seek to minimize the number of integrals in (2-16). We will use the interpolative separable density fitting decomposition (ISDF) [19; 20]. For periodic systems, we first consider the general form of decomposition

$$
Z_{i\boldsymbol{k},j\boldsymbol{k}'}(\boldsymbol{r}) := u_{i\boldsymbol{k}}(\boldsymbol{r})\bar{u}_{j\boldsymbol{k}'}(\boldsymbol{r}) \approx \sum_{\mu=1}^{N_\mu} \zeta_\mu(\boldsymbol{r}) u_{i\boldsymbol{k}}(\hat{\boldsymbol{r}}_\mu)\bar{u}_{j\boldsymbol{k}'}(\hat{\boldsymbol{r}}_\mu). \tag{3-1}
$$

When the unit cell is discretized into a uniform grid $\{\boldsymbol{r}_n\}_{n=1}^{N_g}$, Z can be viewed as a matrix with its row index being $\boldsymbol{r}$, and the column index being a multi-index $(i\boldsymbol{k},j\boldsymbol{k}')$. The matrix size is thus $N_g \times N^2 N_k^2$ (recall that $N = N_v + N_c$). For a given $\boldsymbol{r}$, $u_{i\boldsymbol{k}}(\boldsymbol{r})\bar{u}_{j\boldsymbol{k}'}(\boldsymbol{r})$ can be viewed as a row vector of size $N^2 N_k^2$. The ISDF decomposition then states that all such matrix rows can be approximately expanded using a linear combination of matrix rows with respect to a selected set of *interpolation points* $\{\hat{\boldsymbol{r}}_\mu\}_{\mu=1}^{N_\mu} \subset \{\boldsymbol{r}_i\}_{i=1}^{N_g}$. The coefficients of such a linear combination, or *interpolating vectors*, are denoted by $\{\zeta_\mu(\boldsymbol{r})\}_{\mu=1}^{N_\mu}$. Here N_μ can be interpreted as the numerical rank of the ISDF decomposition.

The compression of the pair products $u_{ik}(r)\bar{u}_{jk'}(r)$ can be understood from the following two limits. First, if only the Γ-point is used to sample the Brillouin zone, we find that there are $N_v N_c \sim N^2$ pairs of functions. However, the number of grid points N_g only scales linearly with respect to N. Hence, the numerical rank of the pair products must scale asymptotically as $\mathbb{O}(N)$. In fact, when all orbitals are smooth functions, we can expect the numerical rank N_μ to be much lower than N_g. This statement has been confirmed by recent analysis [17]. Second, if a large number of k-points are used to discretize the Brillouin zone, N_v, N_c are often relatively small, and the number of grid points in the unit cell N_g does not increase with respect to N_k. Hence, as N_k increases, we may also expect that the numerical rank N_μ will be determined by smoothness of u with respect to r, k, and is asymptotically independent of N_k. This is indeed what has been observed numerically [20]. Throughout the discussion below, we will focus on the second scenario, i.e., we will explicitly write down the scaling with respect to N_g, N, and N_k, but we will primarily focus on the scaling with respect to N_k.

Assuming the interpolation points $\{\hat{r}_\mu\}_{\mu=1}^{N_\mu}$ are already chosen, the interpolation vectors can be efficiently evaluated using a least squares method as follows [12]. Using a linear algebra notation, (3-1) can be written as

$$Z \approx \Theta C. \tag{3-2}$$

Here $\Theta = [\zeta_1, \zeta_2, \ldots, \zeta_{N_\mu}]$ contains the interpolating vectors. Each column of C indexed by (ik, jk') is given by

$$[u_{ik}(\hat{r}_1)\bar{u}_{jk'}(\hat{r}_1), \ldots, u_{ik}(\hat{r}_\mu)\bar{u}_{jk'}(\hat{r}_\mu), \ldots, u_{ik}(\hat{r}_{N_\mu})\bar{u}_{jk'}(\hat{r}_{N_\mu})]^\top.$$

Equation (3-2) is an over-determined linear system with respect to the interpolation vectors Θ. The least squares approximation to the solution is given by

$$\Theta = ZC^*(CC^*)^{-1}. \tag{3-3}$$

Due to the tensor product structure of Z and C, the matrix–matrix multiplications ZC^* and CC^* can be carried out efficiently [12], with computational cost $\mathbb{O}(N_g N_\mu N N_k)$ and $\mathbb{O}(N_\mu^2 N N_k)$, respectively. The cost of inverting the matrix CC^* is $\mathbb{O}(N_\mu^3)$, and the overall cost of evaluating Θ is thus bounded by $\mathbb{O}(N_g N_\mu N N_k + N_\mu^3 + N_g N_\mu^2)$. Hence, the cost scales cubically with respect to the number of electrons in the unit cell, and linearly with respect to the number of k-points.

Equation (3-1) is the general form of ISDF. In the BSE calculations, we may further distinguish whether i, j should take valence or conduction band indices only, as well as whether k, k' can be set to be the same. For instance, (2-17) suggests

that in order to compress V_A, V_B, we only need the ISDF decomposition

$$Z^V_{i_c i_v \boldsymbol{k}}(\boldsymbol{r}) := u_{i_c \boldsymbol{k}}(\boldsymbol{r})\bar{u}_{i_v \boldsymbol{k}}(\boldsymbol{r}) \approx \sum_{\mu=1}^{N^V_\mu} \zeta^V_\mu(\boldsymbol{r}) u_{i_c \boldsymbol{k}}(\hat{\boldsymbol{r}}_\mu)\bar{u}_{i_v \boldsymbol{k}}(\hat{\boldsymbol{r}}_\mu). \qquad (3\text{-}4)$$

Note that the number of columns of the matrix Z^V is only $N_v N_c N_k$, and the number of fitting functions N^V_μ can be chosen to be less than N_μ. The computation of W_A, W_B requires the general ISDF format (3-1).

The interpolation points $\{\hat{\boldsymbol{r}}_\mu\}_{\mu=1}^{N_\mu}$ can be chosen in different ways. In this work we employ a randomized variant of QR with column pivoting (QRCP) [19; 20; 9]. Another recently developed method is based on the centroidal Voronoi decomposition (CVT) [8]. We observed that in our examples it is even possible to work with coarse uniform grids as interpolation points, reducing the computational effort for finding the points to essentially zero while only slightly increasing the error. Since the computation of interpolation points is not the bottleneck in our problem, however, we stick to the previously developed techniques.

4. Fast algorithm for applying the BSH to a vector

Once the ISDF decomposition is obtained, we may compute the matrix elements

$$\widetilde{V}_{A,\mu\nu} = \mathcal{V}(\zeta^V_\mu, \zeta^V_\nu), \quad \widetilde{V}_{B,\mu\nu} = \mathcal{V}(\zeta^V_\mu, \bar{\zeta}^V_\nu), \quad \mu, \nu = 1, \ldots, N^V_\mu, \qquad (4\text{-}1)$$

and similarly

$$\widetilde{W}_{\boldsymbol{q},\mu\nu} = \mathcal{W}_{\boldsymbol{q}}(\zeta_\mu, \zeta_\nu), \quad \mu, \nu = 1, \ldots, N_\mu. \qquad (4\text{-}2)$$

The expressions in (2-17) can then be approximated in the ISDF format as

$$V_A(i_v i_c \boldsymbol{k}, \, j_v j_c \boldsymbol{k}') \approx \frac{1}{N_k} \sum_{\mu,\nu=1}^{N^V_\mu} \bar{u}_{i_c \boldsymbol{k}}(\hat{\boldsymbol{r}}_\mu) u_{i_v \boldsymbol{k}}(\hat{\boldsymbol{r}}_\mu) \widetilde{V}_{A,\mu\nu} \bar{u}_{j_c \boldsymbol{k}'}(\hat{\boldsymbol{r}}_\nu) u_{j_c \boldsymbol{k}'}(\hat{\boldsymbol{r}}_\nu),$$

$$V_B(i_v i_c \boldsymbol{k}, \, j_v j_c \boldsymbol{k}') \approx \frac{1}{N_k} \sum_{\mu,\nu=1}^{N^V_\mu} \bar{u}_{i_c \boldsymbol{k}}(\hat{\boldsymbol{r}}_\mu) u_{i_v \boldsymbol{k}}(\hat{\boldsymbol{r}}_\mu) \widetilde{V}_{B,\mu\nu} \bar{u}_{j_c \boldsymbol{k}'}(\hat{\boldsymbol{r}}_\nu) u_{j_v \boldsymbol{k}'}(\hat{\boldsymbol{r}}_\nu),$$

$$W_A(i_v i_c \boldsymbol{k}, \, j_v j_c \boldsymbol{k}') = \frac{1}{N_k} \sum_{\mu,\nu=1}^{N_\mu} \bar{u}_{i_c \boldsymbol{k}}(\hat{\boldsymbol{r}}_\mu) u_{j_c \boldsymbol{k}'}(\hat{\boldsymbol{r}}_\mu) \widetilde{W}_{\boldsymbol{k}-\boldsymbol{k}',\mu\nu} \bar{u}_{j_v \boldsymbol{k}'}(\hat{\boldsymbol{r}}_\nu) u_{i_v \boldsymbol{k}}(\hat{\boldsymbol{r}}_\nu),$$

$$W_B(i_v i_c \boldsymbol{k}, \, j_v j_c \boldsymbol{k}') = \frac{1}{N_k} \sum_{\mu,\nu=1}^{N_\mu} \bar{u}_{i_c \boldsymbol{k}}(\hat{\boldsymbol{r}}_\mu) u_{j_v \boldsymbol{k}'}(\hat{\boldsymbol{r}}_\mu) \widetilde{W}_{\boldsymbol{k}-\boldsymbol{k}',\mu\nu} \bar{u}_{j_c \boldsymbol{k}'}(\hat{\boldsymbol{r}}_\nu) u_{i_v \boldsymbol{k}}(\hat{\boldsymbol{r}}_\nu).$$

$(4\text{-}3)$

In order to use the Fourier representation (2-33) and (2-36), we first need to per-form Fourier transforms for $\{\zeta_\mu^V\}$ and $\{\zeta_\mu\}$. Using the fast Fourier transform (FFT), and assuming that the number of Fourier coefficients G is also N_g, the computational cost for the Fourier transform scales as $\mathbb{O}(N_\mu^V N_g \log N_g)$ and $\mathbb{O}(N_\mu N_g \log N_g)$, respectively. The Fourier coefficients $\widehat{V}_k$ can be obtained analytically, and we assume the coefficients $\widehat{W}_k$ are already provided from, e.g., a GW calculation. The cost for computing $\widetilde{V}_A$, $\widetilde{V}_B$ using (2-33) is then $\mathbb{O}((N_\mu^V)^2 N_g)$. Similarly the cost for computing all $\widetilde{W}_q$ matrices is $\mathbb{O}(N_\mu^2 N_g N_k)$. In particular, the total cost for the initial setup stage scales as $\mathbb{O}(N_k)$ with respect to the number of k-points.

After this initial setup stage, each entry of the BSH can be computed with $\mathbb{O}((N_\mu^V)^2 + N_\mu^2)$ operations. If the entire BSH matrix is to be constructed, the cost will be $\mathbb{O}(N_\mu^2 N_k^2 N_v^2 N_c^2)$.

Below we demonstrate that if we only aim to apply the Hamiltonian H_{BSE} to an arbitrary vector without ever assembling the full Hamiltonian, the computational cost can be greatly reduced.

For simplicity, let us focus on the case when the Tamm–Dancoff approximation (TDA) is used. Applying the Hamiltonian $H_{\text{BSE}} = D + 2V_A - W_B$ to a vector $X \in \mathbb{C}^{N_v N_c N_k}$ amounts to evaluating the three terms

$$
\begin{aligned}
[DX](i_v i_c \mathbf{k}) &= (\epsilon_{i_c \mathbf{k}} - \epsilon_{i_v \mathbf{k}'}) X(i_v i_c \mathbf{k}), \\
[V_A X](i_v i_c \mathbf{k}) &= \sum_{j_v, j_c, \mathbf{k}'} V_A(i_v i_c \mathbf{k}, j_v j_c \mathbf{k}') X(j_v j_c \mathbf{k}'), \\
[W_A X](i_v i_c \mathbf{k}) &= \sum_{j_v, j_c, \mathbf{k}'} W_A(i_v i_c \mathbf{k}, j_v j_c \mathbf{k}') X(j_v j_c \mathbf{k}').
\end{aligned}
\tag{4-4}
$$

Computing the first term for all $(i_v i_c \mathbf{k})$ clearly costs $\mathbb{O}(N_v N_c N_k)$ operations. We now show that the second and third terms can also be computed efficiently.

Using (4-3), the second term in (4-4) can be regrouped as

$$
\frac{1}{N_k} \sum_\mu \bar{u}_{i_c \mathbf{k}}(\hat{\mathbf{r}}_\mu) u_{i_v \mathbf{k}}(\hat{\mathbf{r}}_\mu) \left\{ \sum_\nu \widetilde{V}_{A,\mu\nu} \right.
$$

$$
\left. \times \left(\sum_{\mathbf{k}'} \left(\sum_{j_c} u_{j_c \mathbf{k}'}(\hat{\mathbf{r}}_\nu) \left(\sum_{j_v} \bar{u}_{j_v \mathbf{k}'}(\hat{\mathbf{r}}_\nu) X(j_v j_c \mathbf{k}') \right) \right) \right) \right\}.
\tag{4-5}
$$

This means one can first perform contractions over j_v, j_c, and $\mathbf{k}'$ to obtain a quantity that only depends on $\hat{\mathbf{r}}_\nu$. The computational complexity is $\mathbb{O}(N_\mu^V(N_v N_c N_k + N_c N_k))$. The two remaining sums can be computed with $\mathbb{O}((N_\mu^V)^2 + N_\mu^V N_v N_c N_k)$ operations. The total complexity of computing $V_A X$ is bounded by $\mathbb{O}((N_\mu^V)^2 + N_\mu^V N_v N_c N_k)$.

For the third term in (4-4) we obtain

$$
\frac{1}{N_k} \sum_{v} u_{i_v \boldsymbol{k}}(\hat{\boldsymbol{r}}_v) \left\{ \sum_{\mu} \bar{u}_{i_c \boldsymbol{k}}(\hat{\boldsymbol{r}}_\mu) \right.
$$

$$
\left. \times \left(\sum_{\boldsymbol{k}'} \widetilde{W}_{\boldsymbol{k}-\boldsymbol{k}',\mu v} \left(\sum_{j_c} u_{j_c \boldsymbol{k}'}(\hat{\boldsymbol{r}}_\mu) \left(\sum_{j_v} \bar{u}_{j_v \boldsymbol{k}'}(\hat{\boldsymbol{r}}_v) X(j_v j_c \boldsymbol{k}') \right) \right) \right) \right\}. \quad (4\text{-}6)
$$

Here, we exploited the separable structure of the decomposition to reorder the products in such a way that all terms depending on $\boldsymbol{k}$ and $\boldsymbol{k}'$ are to the left and right, respectively, of $\widetilde{W}_{\boldsymbol{k}-\boldsymbol{k}',\mu v}$. The two innermost contractions over j_v and j_c result in a quantity that only depends on $\boldsymbol{k}$, $\hat{\boldsymbol{r}}_\mu$, and $\hat{\boldsymbol{r}}_v$. The cost for these two steps is $\mathbb{O}(N_\mu N_k N_v N_c + N_\mu^2 N_k N_c)$. The sum over $\boldsymbol{k}'$ then has the structure of a *discrete convolution*, for each fixed μv pair. Therefore, it can be computed for all $\boldsymbol{k}$ simultaneously in $\mathbb{O}(N_\mu^2 N_k \log N_k)$ operations by fast convolution algorithms, e.g., by using the FFT with zero-padded vectors. The remaining summation operations over μ and v are then obtained with $\mathbb{O}(N_\mu^2 N_c N_k + N_\mu N_v N_c N_k)$ operations. In total the computation of $W_A X$ amounts to $\mathbb{O}(N_\mu N_v N_c N_k + N_\mu^2 N_c N_k + N_\mu^2 N_k \log N_k)$ operations.

Combining the results for the three parts of the Hamiltonian, we see that the computational complexity is given by

$$
\mathbb{O}\left((N_\mu + N_\mu^V) N_v N_c N_k + (N_\mu^V)^2 + N_\mu^2 N_c N_k + N_\mu^2 N_k \log N_k \right).
$$

In particular, the cost with respect to the number of $\boldsymbol{k}$-points only scales as $\mathbb{O}(N_k \log N_k)$. This allows us to perform BSE calculations for complex materials which require a very large number of $\boldsymbol{k}$-points.

By avoiding the explicit construction of H_{BSE}, the new algorithm also drastically reduces the storage cost. The storage cost for H_{BSE} alone is $\mathbb{O}((N_v N_c N_k)^2)$. In the new algorithm, the storage cost of $\widehat{W}_{\boldsymbol{q}}$ becomes the dominant component and scales only linearly with respect to N_k.

As an example, the matrix-free application of H_{BSE} can be used to compute the optical absorption spectrum, which requires the evaluation of the quantity

$$
\varepsilon_2(\omega) = \mathrm{Im}\left[\frac{8\pi}{|\Omega|} d_r^* ((\omega - i\eta) I - H_{\mathrm{BSE}})^{-1} d_l \right]. \quad (4\text{-}7)
$$

Here d_r and d_l are called the right and left optical transition vectors, and η is a broadening factor used to account for the exciton lifetime. We also compute the smallest eigenvalues of H_{BSE}, which are of interest in their own right, as they represent the transition energies of bound excitons in many semiconducting solid state materials.

To observe the absorption spectrum and identify its main peaks, it is possible to use a structure-preserving iterative method instead of explicitly computing all eigenpairs of H_{BSE}. We refer readers to [6; 34] for details of the structure-preserving Lanczos algorithm, which has been implemented in the BSEPACK [35] library.[1] When TDA is used, the structure-preserving Lanczos reduces to a standard Lanczos algorithm. For the computation of the first eigenvalue we use standard ARPACK [14] routines for Hermitian matrices.

5. Numerical examples

To illustrate the efficiency of ISDF for BSE calculations in crystals, we apply the method to compute the excitation modes and absorption spectra of a one-dimensional model problem as well as two real material systems, diamond (3D bulk) and graphene (quasi-2D). For both systems, we determine the optical absorption spectra on k-grids close to those employed in previously published calculations to demonstrate that our method is suitable for state-of-the-art calculations, both for 3D and quasi-2D materials. We furthermore provide a numerical scaling analysis and a more detailed analysis of the error in the ISDF in the case of the one-dimensional model and diamond. We show that a good approximation of the spectrum can be obtained with a small number of interpolation vectors.

The method was implemented in the programming language Julia [5] and the source code is available.[2] As the input to our method for the actual materials, we employ the KSDFT single particle orbitals, quasiparticle energies, and screened Coulomb potential computed by exciting [10; 37], an all-electron full-potential code with implementations of density functional theory and many-body perturbation theory. The Tamm–Dancoff approximation is used in all calculations.

All calculation for the proposed method were carried out on a single core of an Intel Core i5-8250U CPU at 1.60 GHz.

5.1. *One-dimensional problems.* For the one-dimensional problem, we take the single particle orbitals $\psi_{ik}(r)$ in (2-16) to be eigenfunctions of a single particle Hamiltonian $\mathcal{H}(k)$ in which the effective potential is defined as

$$V_{\text{eff}}(r) = 20\cos(4\pi r/L) + 0.2\sin(2\pi r/L),$$

where the unit cell size is $|\Omega| \equiv L = 1.5$.

The bare Coulomb potential used in (2-16) is chosen to be

$$V(r, r') = \frac{1}{\sqrt{(r-r')^2 + 0.01}}, \tag{5-1}$$

---

[1] https://sites.google.com/a/lbl.gov/bsepack/
[2] https://github.com/fhenneke/BSE_k_ISDF.jl/
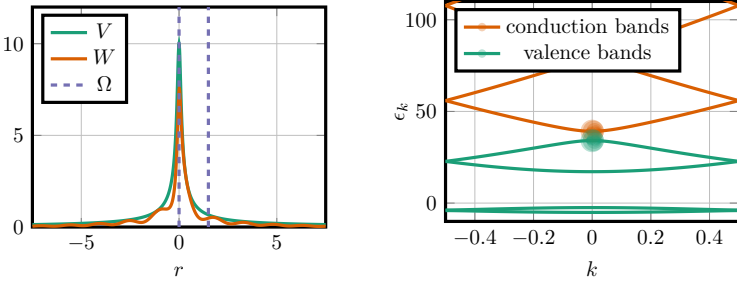
Figure 1. Left: the potentials $V(r, 0)$ and $W(r, 0)$. Right: band structure with coefficients of the lowest eigenfunction for $N_k = 128$. The areas of the circles on the valence and conduction bands at position $\boldsymbol{k}$ are proportional to $\sum_{i_c} |X(i_v i_c \boldsymbol{k})|^2$ and $\sum_{i_v} |X(i_v i_c \boldsymbol{k})|^2$.

and the screened interaction is chosen as

$$W(r, r') = \frac{(3 + \sin(2\pi r/L))(3 + \cos(4\pi r'/L))}{16} e^{-(r-r')^2/(32L^2)} V(r, r'). \quad (5\text{-}2)$$

Compared to the smoothed-out Coulomb potential V, the chosen screened interaction W decays exponentially and also contains lattice periodic contributions. The potentials are shown in Figure 1. Both potentials are periodically extended $N_k - 1$ times outside of the unit cell. The particular structure of the potentials has an influence on the band structure and spectrum of the BSH, but was observed to not significantly impact the convergence behavior or the run time scaling of the ISDF method.

The Bloch functions u_{ik} are sampled on $N_g = 128$ uniformly distributed grid points within the unit cell, and the number of $\boldsymbol{k}$-points N_k ranges from 16 to 4096 in our experiments.

For each $\boldsymbol{k}$-point, the first four eigenstates are treated as the valence states in this model, while the remaining eigenstates are considered as the conduction states, separated by an energy gap from the former. We use all $N_v = 4$ valence bands and $N_c = 5$ conduction bands to construct the approximate H_{BSE}. The number of $\boldsymbol{k}$-points was chosen to be $N_k = 256$ in the error analysis of the ISDF approximation, and varies from 16 to 4096 in the run time analysis and the analysis of the error in the absorption spectrum. The largest resulting Hamiltonian is of size $81\,920 \times 81\,920$.

Figure 2 shows how the ISDF approximation error varies with respect to the truncation parameter N_μ^{ij} and how the accuracy of the approximate spectrum of H_{BSE} changes with respect to the ISDF approximation error.

In the left subfigure, we plot the relative error $\|\Theta^{\alpha\beta} C^{\alpha\beta} - Z^{\alpha\beta}\|_F / \|Z^{\alpha\beta}\|_F$, $\alpha, \beta \in \{v, c\}$, where $\|\cdot\|_F$ is the Frobenius norm, for different choices of truncation levels N_μ (or number of interpolation points). As expected, when N_μ is too small, ISDF results in relatively large error. As N_μ becomes slightly larger, the ISDF approximation error decays exponentially with respect to N_μ up to $N_\mu = 20 \sim 30$.
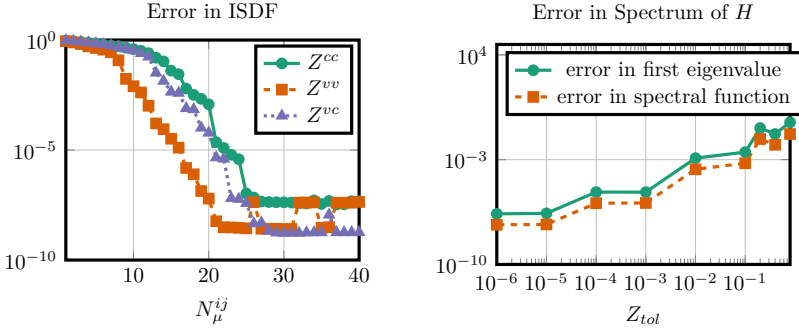
Figure 2. Left: ISDF approximation error $\|Z - \Theta C\|_F / \|Z\|_F$ for different choices of N_μ. Right: resulting errors in the spectrum of H_{BSE} for different ISDF error tolerances.
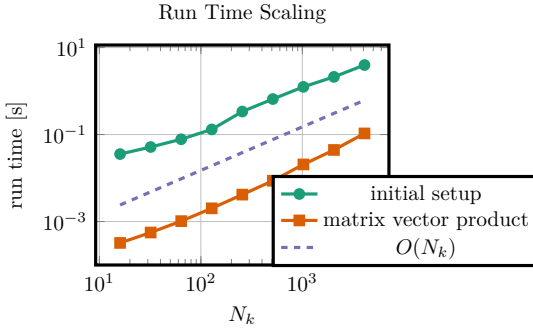
Figure 3. Run times for the initial setup and individual matrix-free matrix-vector products.

At this truncation level, the error is on the order of 10^{-8}, which is sufficiently small for obtaining a highly accurate approximation of the spectrum of H_{BSE} as shown in the right subfigure. In this subfigure, we plot the relative error in the first eigenvalue and in the overall optical absorption spectrum against the ISDF error tolerance Z_{tol}. For each Z_{tol}, we choose the smallest truncation parameters N_μ with the resulting error in $Z^{\alpha,\beta}$ being less than or equal to Z_{tol} for $\alpha, \beta \in \{v, c\}$.

In Figure 3, we plot the timing measurements for both the construction of $\widetilde{V}$ and $\widetilde{W}$ and the multiplication of the approximate H_{BSE} with a vector with respect to N_k. In these calculations, the ISDF truncation parameters N_μ are chosen so that the relative error in $Z^{\alpha\beta}$ is below $Z_{\mathrm{tol}} = 10^{-5}$. This error tolerance resulted in the choices of $N_\mu^{vv} = 17$, $N_\mu^{cc} = 23$, and $N_\mu^{vc} = 21$.

As we can see in Figure 3, the scaling of the run time for the construction of $\widetilde{V}$ and $\widetilde{W}$ is nearly linear with respect to N_k, which is in excellent agreement with the theoretical computational complexity presented in the preceding section. The scaling of the run time for the multiplication of the approximate H_{BSE} with a vector also looks linear in N_k. In fact, a more detailed investigation showed that the

Optical Absorption of Diamond

Optical Absorption of Graphene
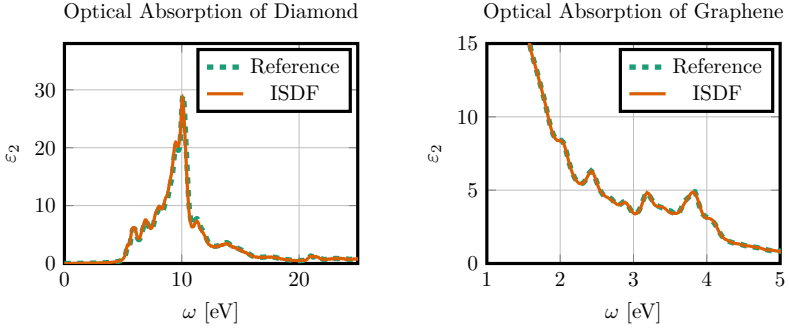
Figure 4. Optical absorption spectrum for diamond (left) and graphene (right).

parameters	diamond	graphene
N_v	4	4
N_c	10	5
N_k	$13 \times 13 \times 13$	$42 \times 42 \times 1$
N_r	$20 \times 20 \times 20$	$15 \times 15 \times 50$
N_μ^{vv}	70	50
N_μ^{cc}	220	180
N_μ^{vc}	100	60
N_{iter}	150	100

Table 1. Parameters used in the computation of spectra and the benchmarks.

convolutions in k in the application of W dominate the cost of the matrix-vector multiplications, in good agreement with the theoretical $\mathbb{O}(N_k \log N_k)$ complexity shown earlier.

For comparison, without the use of ISDF, the construction of H_{BSE} is estimated to take about 460 000 seconds for $N_k = 4096$. With our method it took less than 10 seconds.

5.2. Three-dimensional problems. We now compare optical absorption spectra for diamond and graphene computed from the approximate H_{BSE} constructed via ISDF with corresponding reference spectra. The reference spectra are obtained from the exact H_{BSE} from the `exciting` code [10; 37]. The comparison is shown in Figure 4. The reference spectrum for diamond is constructed on a $13 \times 13 \times 13$ k-grid using all 4 valence and 10 conduction states. Fourier components $\widehat{W}_q(G, G')$ in (2-35) are calculated up to a cutoff $|G + q| \leq 2.5 \, a_0^{-1}$, where a_0 is the Bohr radius. The screened Coulomb interaction is calculated within the random-phase approximation (RPA) including 100 conduction states. For graphene, the reference spectrum is obtained on a $42 \times 42 \times 1$ k-grid using all 4 valence and 5 conduction states. Fourier
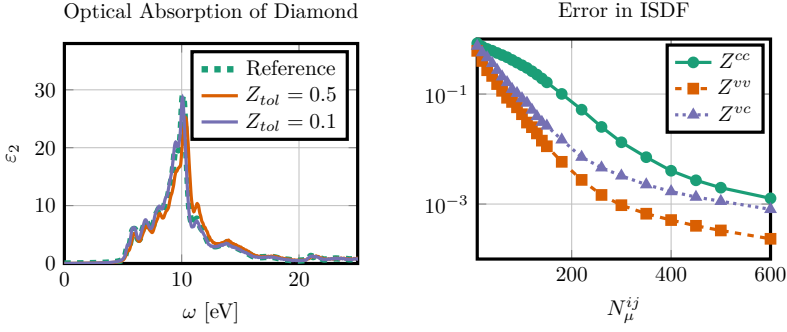
Figure 5. Left: optical absorption spectrum for diamond with differently accurate ISDF approximations. Right: estimated errors in ISDF approximation with different numbers of interpolation points.

Z_{tol}	error in		
	absorption function	first eigenvalue	
0.5	0.199	0.0038	(20.7 meV)
0.1	0.056	0.0011	(6.2 meV)
0.05	0.040	0.0006	(3.3 meV)

Table 2. Relative (and absolute) errors in the spectrum of H_{BSE} for different ISDF error tolerances.

components $\widehat{W}_{\boldsymbol{q}}(\boldsymbol{G}, \boldsymbol{G}')$ in (2-35) are calculated up to a cutoff $|\boldsymbol{G} + \boldsymbol{q}| \leq 2.0\,\mathrm{a}_0^{-1}$, and 80 conduction states are included in the RPA calculations for the screened Coulomb potential. The numerical parameters of the reference and approximate calculations are shown in Table 1. The number of interpolation vectors was chosen such that the relative ISDF error was around 0.1.

We can clearly see that for both diamond and graphene, the approximate optical absorption spectrum matches well with the reference spectrum. In particular, the positions and heights of all major peaks are in good agreement. We should note that, in the case of diamond, the absorption spectrum produced by a $13 \times 13 \times 13$ $\boldsymbol{k}$-grid is in good agreement with measurements [26] and previous BSE calculations [11]. In the case of graphene, however, larger $\boldsymbol{k}$-grids have been reported for BSE calculations [38] to produce an optical absorption spectrum in good agreement with the experimental result.

Figure 5 shows that the ISDF approximation error can be systematically reduced as we increase the number interpolating vectors N_μ. However, Figure 4 shows that the approximate absorption spectrum is already in good agreement with the reference spectrum, when the relative ISDF approximation error is at 0.1. Thus, it seems unnecessary to use a larger number of interpolation vectors in these cases. This observation is corroborated by the relative difference between the first eigenvalue
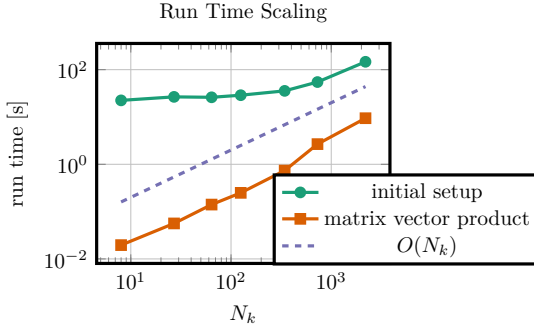
Figure 6. Run times for the initial setup and individual matrix-free matrix-vector products.

of the approximate H_{BSE} computed using ARPACK and that of reference H_{BSE} constructed in `exciting` shown in Table 2. With a relative ISDF approximation error of $Z_{\text{tol}} = 0.1$, the error in the first BSE eigenvalue is below $10\,\text{meV}$ in both examples shown here.

To illustrate the run time scaling of the method in the 3D examples, we measure the time it takes to construct the approximate H_{BSE} via ISDF as well as the time it takes to multiply the resulting H_{BSE} with vectors for the diamond example. We use $\boldsymbol{k}$-grids of sizes $N_k = n_k \times n_k \times n_k$ for $n_k \in \{2, 3, 4, 5, 7, 9, 13\}$. The resulting timing measurements are plotted in Figure 6. It can be seen that the run time for constructing the approximate H_{BSE} scales linearly with the number of $\boldsymbol{k}$-points. The multiplication of H_{BSE} with vectors scales as $\mathbb{O}(N_k \log N_k)$ for sufficiently large N_k. As in the model problem, the convolutions in $\boldsymbol{k}$ in the application of W dominate the cost of the matrix–vector multiplications. For comparison, computing the ISDF decomposition of the Hamiltonian for the case $N_k = 13^3$ took 147 seconds, whereas the full assembly of the Hamiltonian took about 6 hours in `exciting` on 13 compute nodes with 13 cores each. The optical absorption function was obtained by running about 150 Lanczos steps, which amounts to about 24 minutes for each fixed direction (x, y, and z), compared to almost 4 hours required in the `exciting` code for the full diagonalization on 13 compute nodes.

6. Conclusion

In this paper, we examined the possibility of using the ISDF technique to reduce the computational complexity of BSH construction and the subsequent iterative approximation of the optical absorption spectrum and excitation energies of electron-hole (exciton) pairs for solids. For periodic systems, a fine $\boldsymbol{k}$-point sampling in the Brillouin zone is often required to produce accurate results, whereas the number of bands per $\boldsymbol{k}$-point required to construct the bare exchange and screened direct kernels of the BSH is relatively small. We showed that the complexity of the ISDF

procedure scales linearly with respect to the number of $\boldsymbol{k}$-points (N_k) when the ranks of the approximate bare exchange and screened direct kernels produced by the ISDF procedure are chosen to be independent of N_k. By keeping the bare exchange and screened direct kernels in the low-rank decomposed form produced by the ISDF procedure, an iterative method used to obtain the optical absorption spectrum and selected excitation energies (eigenvalues of the BSH) can be implemented with cost scaling as $\mathbb{O}(N_k \log N_k)$. Our numerical experiments, which were performed on a 1D model as well as two different types of actual materials (diamond and graphene), confirm our complexity analysis. They demonstrate that the ISDF technique can indeed significantly reduce the cost of BSE calculation for solids while maintaining the same accuracy provided by a standard BSE calculation implemented in the software `exciting`. Our current implementation of the ISDF technique is done using the Julia programming language for a single node. A distributed parallel implementation is needed to accommodate a much finer $\boldsymbol{k}$-point sampling which is required in the case of the graphene example to produce a computed absorption spectrum that matches with experimental results.

Acknowledgments

References

[1] S. Albrecht, G. Onida, and L. Reining, *Ab initio calculation of the quasiparticle spectrum and excitonic effects in* Li_2O, Phys. Rev. B **55** (1997), no. 16, 10278–10281.

[2] N. W. Ashcroft and N. D. Mermin, *Solid state physics*, Harcourt, New York, 1976.

[3] P. Benner, V. Khoromskaia, and B. N. Khoromskij, *A reduced basis approach for calculation of the Bethe–Salpeter excitation energies by using low-rank tensor factorisations*, Mol. Phys. **114** (2016), no. 7–8, 1148–1161.

[4] P. Benner, S. Dolgov, V. Khoromskaia, and B. N. Khoromskij, *Fast iterative solution of the Bethe–Salpeter eigenvalue problem using low-rank and QTT tensor approximation*, J. Comput. Phys. **334** (2017), 221–239. MR Zbl

[5] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, *Julia: a fresh approach to numerical computing*, SIAM Rev. **59** (2017), no. 1, 65–98. MR Zbl

[6] 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*, J. Chem. Theory Comput. **11** (2015), no. 11, 5197–5208.

[7] J. Deslippe, G. Samsonidze, D. A. Strubbe, M. Jain, M. L. Cohen, and S. G. Louie, *BerkeleyGW: a massively parallel computer package for the calculation of the quasiparticle and optical properties of materials and nanostructures*, Comput. Phys. Commun. **183** (2012), no. 6, 1269–1289.

[8] K. Dong, W. Hu, and L. Lin, *Interpolative separable density fitting through centroidal Voronoi tessellation with applications to hybrid functional electronic structure calculations*, J. Chem. Theory Comput. **14** (2018), no. 3, 1311–1320.

[9] G. H. Golub and C. F. Van Loan, *Matrix computations*, 4th ed., Johns Hopkins University, Baltimore, MD, 2013. MR Zbl

[10] A. Gulans, S. Kontur, C. Meisenbichler, D. Nabok, P. Pavone, S. Rigamonti, S. Sagmeister, U. Werner, and C. Draxl, `exciting`: *a full-potential all-electron package implementing density-functional theory and many-body perturbation theory*, J. Phys. Condens. Mat. **26** (2014), no. 36, art. id. 363202.

[11] P. H. Hahn, K. Seino, W. G. Schmidt, J. Furthmüller, and F. Bechstedt, *Quasiparticle and excitonic effects in the optical spectra of diamond,* SiC, Si, GaP, GaAs, InP, and AlN, Phys. Status Solidi B **242** (2005), no. 13, 2720–2728.

[12] W. Hu, L. Lin, and C. Yang, *Interpolative separable density fitting decomposition for accelerating hybrid density functional calculations with applications to defects in silicon*, J. Chem. Theory Comput. **13** (2017), no. 11, 5420–5431.

[13] W. Hu, M. Shao, A. Cepellotti, F. H. da Jornada, L. Lin, K. Thicke, C. Yang, and S. G. Louie, *Accelerating optical absorption spectra and exciton energy computation via interpolative separable density fitting*, ICCS 2018, II (Y. Shi, H. Fu, Y. Tian, V. V. Krzhizhanovskaya, M. H. Lees, J. Dongarra, and P. M. A. Sloot, eds.), Lecture Notes in Comput. Sci., no. 10861, Springer, 2018, pp. 604–617. MR

[14] R. B. Lehoucq, D. C. Sorensen, and C. Yang, *ARPACK users' guide: solution of large-scale eigenvalue problems with implicitly restarted arnoldi methods*, Software, Environments, and Tools, no. 6, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1998. MR Zbl

[15] L. Lin, Z. Xu, and L. Ying, *Adaptively compressed polarizability operator for accelerating large scale ab initio phonon calculations*, Multiscale Model. Simul. **15** (2017), no. 1, 29–55. MR Zbl

[16] M. P. Ljungberg, P. Koval, F. Ferrari, D. Foerster, and D. Sánchez-Portal, *Cubic-scaling iterative solution of the Bethe–Salpeter equation for finite systems*, Phys. Rev. B **92** (2015), no. 7, art. id. 075422.

[17] J. Lu, C. D. Sogge, and S. Steinerberger, *Approximating pointwise products of Laplacian eigenfunctions*, J. Funct. Anal. **277** (2019), no. 9, 3271–3282. MR Zbl

[18] J. Lu and K. Thicke, *Cubic scaling algorithms for RPA correlation using interpolative separable density fitting*, J. Comput. Phys. **351** (2017), 187–202. MR Zbl

[19] J. Lu and L. Ying, *Compression of the electron repulsion integral tensor in tensor hypercontraction format with cubic scaling cost*, J. Comput. Phys. **302** (2015), 329–335. MR Zbl

[20] _____ , *Fast algorithm for periodic density fitting for Bloch waves*, Ann. Math. Sci. Appl. **1** (2016), no. 2, 321–339. MR Zbl

[21] M. Marsili, F. Mosconi, Edoardo De Angelis, and P. Umari, *Large-scale GW-BSE calculations with N^3 scaling: excitonic effects in dye-sensitized solar cells*, Phys. Rev. B **95** (2017), no. 7, art. id. 075415.

[22] H. J. Monkhorst and J. D. Pack, *Special points for Brillouin-zone integrations*, Phys. Rev. B **13** (1976), no. 12, 5188–5192. MR

[23] N. L. Nguyen, H. Ma, M. Govoni, F. Gygi, and G. Galli, *Finite-field approach to solving the Bethe–Salpeter equation*, Phys. Rev. Lett. **122** (2019), no. 23, art. id. 237402. MR

[24] G. Onida, L. Reining, R. W. Godby, R. Del Sole, and W. Andreoni, *Ab initio calculations of the quasiparticle and absorption spectra of clusters: the sodium tetramer*, Phys. Rev. Lett. **75** (1995), no. 5, 818–821.

[25] G. Onida, L. Reining, and A. Rubio, *Electronic excitations: density-functional versus many-body Green's-function approaches*, Rev. Mod. Phys. **74** (2002), no. 2, 601–659.

[26] H. R. Phillip and E. A. Taft, *Kramers–Kronig analysis of reflectance data for diamond*, Phys. Rev. **136** (1964), no. 5A, A1445–A1448.

[27] Y. Ping, D. Rocca, and G. Galli, *Electronic excitations in light absorbers for photoelectrochemical energy conversion: first principles calculations based on many body perturbation theory*, Chem. Soc. Rev. **42** (2013), 2437–2469.

[28] Y. Ping, D. Rocca, D. Lu, and G. Galli, *Ab initio calculations of absorption spectra of semiconducting nanowires within many-body perturbation theory*, Phys. Rev. B **85** (2012), no. 3, art. id. 035316.

[29] D. Y. Qiu, F. H. da Jornada, and S. G. Louie, *Optical spectrum of MoS_2: many-body effects and diversity of exciton states*, Phys. Rev. Lett. **111** (2013), no. 21, art. id. 216805.

[30] D. Rocca, D. Lu, and G. Galli, *Ab initio calculations of optical absorption spectra: solution of the Bethe–Salpeter equation within density matrix perturbation theory*, J. Chem. Phys. **133** (2010), no. 16, art. id. 164109.

[31] D. Rocca, Y. Ping, R. Gebauer, and G. Galli, *Solution of the Bethe–Salpeter equation without empty electronic states: application to the absorption spectra of bulk systems*, Phys. Rev. B **85** (2012), no. 4, art. id. 045116.

[32] M. Rohlfing and S. G. Louie, *Electron-hole excitations and optical spectra from first principles*, Phys. Rev. B **62** (2000), no. 8, 4927–4944.

[33] E. E. Salpeter and H. A. Bethe, *A relativistic equation for bound-state problems*, Phys. Rev. **84** (1951), 1232–1242. MR Zbl

[34] M. Shao, F. H. da Jornada, L. Lin, C. Yang, J. Deslippe, and S. G. Louie, *A structure preserving Lanczos algorithm for computing the optical absorption spectrum*, SIAM J. Matrix Anal. Appl. **39** (2018), no. 2, 683–711. MR Zbl

[35] M. Shao and C. Yang, *BSEPACK user's guide*, user manual, 2016. arXiv

[36] G. Strinati, *Application of the Green's functions method to the study of the optical properties of semiconductors*, Riv. Nuovo Cimento **11** (1988), no. 12, 1–86.

[37] C. Vorwerk, B. Aurich, C. Cocchi, and C. Draxl, *Bethe–Salpeter equation for absorption and scattering spectroscopy: implementation in the `exciting` code*, Electron. Struct. **1** (2019), no. 3, art. id. 037001.

[38] L. Yang, J. Deslippe, C.-H. Park, M. L. Cohen, and S. G. Louie, *Excitonic effects on the optical response of graphene and bilayer graphene*, Phys. Rev. Lett. **103** (2009), no. 18, art. id. 186802.

FELIX HENNEKE: felix.henneke@fu-berlin.de
Institut für Mathematik, Freie Universität Berlin, Berlin, Germany

LIN LIN: linlin@math.berkeley.edu
Department of Mathematics, University of California, Berkeley, Berkeley, CA, United States

and

Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, CA, United States

CHRISTIAN VORWERK: vorwerk@physik.hu-berlin.de
Institut für Physik, IRIS Adlershof, Humboldt-Universität zu Berlin, Berlin, Germany

CLAUDIA DRAXL: claudia.draxl@physik.hu-berlin.de
Institut für Physik, IRIS Adlershof, Humboldt-Universität zu Berlin, Berlin, Germany

RUPERT KLEIN: rupert.klein@fu-berlin.de
Institut für Mathematik, Freie Universität Berlin, Berlin, Germany

CHAO YANG: cyang@lbl.gov
Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, CA, United States

Communications in Applied Mathematics and Computational Science

msp.org/camcos

See inside back cover or msp.org/camcos for submission instructions.