The Journal of Open Source Software

Low scaling BSE implementation in the exciting code

Benedikt Maurer [1*] **and Claudia Draxl** [1*]

1 Department of Physics and CSMB Adlershof, Humboldt-Universität zu Berlin, Zum Großen Windkanal 2, D-12489 Berlin, Germany * These authors contributed equally.

DOI: 10.21105/joss.08866

Software

- Review
- Repository
- Archive

Editor: Bonan Zhu

Reviewers:

- @jjkas
- @ruiyiQM
- @EderGio

Submitted: 03 February 2025
Published: 03 February 2026

Summary

Solving the Bethe-Salpeter Equation (BSE) is essential for understanding excited-state systems, but often challenging to converge or even computationally prohibitive for large systems. We implement a matrix-free BSE solver leveraging Interpolative Separable Density Fitting (ISDF) to interpolate electron-hole interaction kernels together with the Lanczos algorithm for diagonalization, avoiding full matrix setup. The scaling of our implementation is bounded by $\mathcal{O}(N_o N_u N_\mathbf{k} \log N_\mathbf{k})$ and is thus a massive improvement over methods that set up the whole matrix, scaling with at least $\mathcal{O}((N_o N_u N_\mathbf{k})^3)$, where N_o and N_u are the numbers of occupied and unoccupied states, respectively, and $N_\mathbf{k}$ is the number $\mathbf{k}$-points.

Theoretical background

The Bethe-Salpeter Equation (BSE) within many body perturbation theory (MBPT) provides the state-of-the-art framework for describing light-matter interaction. In particular, it is used to obtain optical absorption spectra, including the effects of excitons, which are bound electron-hole states. By expanding the electron-hole wavefunctions in the transition basis, solving the BSE can be reduced to a Schrödinger-like equation. Setting up and diagonalizing the Bethe-Salpeter Hamiltonian (BSH) are the computationally expensive tasks (Vorwerk et al., 2019). The BSH is given as

$$H^{BSH} = D + \gamma V - W\,,$$

where $\gamma = 2$ gives the spin-singlet and $\gamma = 0$ spin-triplet channel, respectively. The diagonal term D is given by the differences of the one-particle energies of the occupied (o) and unoccupied (u) states:

$$D_{o\mathbf{u}\mathbf{k},o'u'\mathbf{k}'} = (\varepsilon_{u\mathbf{k}} - \varepsilon_{o\mathbf{k}})\delta_{oo'}\delta_{uu'}\delta_{\mathbf{k}\mathbf{k}'}\,.$$

The matrix elements of the repulsive exchange interaction V and the attractive screened Coulomb interaction W are calculated by solving integrals of the form

$$V_{o\mathbf{u}\mathbf{k},o'u'\mathbf{k}'} = \int d^3r \int d^3r' \frac{u_{o\mathbf{k}}(\mathbf{r})u^*_{u\mathbf{k}}(\mathbf{r})u^*_{o'\mathbf{k}'}(\mathbf{r}')u_{u'\mathbf{k}'}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,,$$

$$W_{o\mathbf{u}\mathbf{k},o'u'\mathbf{k}'} = \int d^3r \int d^3r' u^*_{u\mathbf{k}}(\mathbf{r})u_{u'\mathbf{k}'}(\mathbf{r})W_{\mathbf{k}-\mathbf{k}'}(\mathbf{r},\mathbf{r}')u_{o\mathbf{k}}(\mathbf{r}')u^*_{o'\mathbf{k}'}(\mathbf{r}')\,,$$

where $W_{\mathbf{k}-\mathbf{k}'}(\mathbf{r},\mathbf{r}')$ is the statically screened Coulomb potential. $u_{i\mathbf{k}}(\mathbf{r})$ is the periodic part of the one-particle wavefunction of state i at the reciprocal lattice point $\mathbf{k}$. In general, we cannot see which of the matrix elements will be zero, thus we need to compute all of them. Therefore, setting up the full BSH scales with $\mathcal{O}(N_o^2 N_u^2 N_\mathbf{k}^2)$ and diagonalizing it directly with $\mathcal{O}(N_o^3 N_u^3 N_\mathbf{k}^3)$.

The primary scaling bottleneck stems from evaluating the matrix elements of the interaction kernels. To mitigate this, we reformulate the wavefunction products using Interpolative Separable Density Fitting (ISDF) (Lu & Ying, 2015, 2016). Specifically, we approximate these products on a discrete real-space grid, $\{\mathbf{r}\}$, by expressing them as superpositions of values evaluated on a smaller interpolation grid, $\{\mathbf{r}_\mu\} \subset \{\mathbf{r}\}$:

$$u^*_{i\mathbf{k}}(\mathbf{r})u_{j\mathbf{k}'}(\mathbf{r}) \approx \sum_{\mu=1}^{N_\mu} \zeta_\mu(\mathbf{r})u^*_{i\mathbf{k}}(\mathbf{r}_\mu)u_{j\mathbf{k}}(\mathbf{r}_\mu) \,,$$

where N_μ is the number of interpolation points, and $\zeta_\mu(\mathbf{r})$ are the expansion coefficients. Due to the tensor product structure of $u^*_{i\mathbf{k}}(\mathbf{r})u_{j\mathbf{k}'}(\mathbf{r})$, $\zeta_\mu(\mathbf{r})$ can be computed efficiently, and the scaling is bounded by $\mathcal{O}(N_\mu^3)$ (Hu et al., 2017). We observe that we can always choose $N_\mu \ll N_o N_u N_\mathbf{k}$, thus ISDF is never a bottleneck. The interpolation points are computed efficiently with centroidal Voronoi tessellation within $\mathcal{O}(N_\mu N_r)$ (Dong et al., 2018).

Replacing the wavefunction products in the interaction kernel integrals with the ISDF representation yields a formulation that allows for efficient application to a vector without ever setting up the full matrix. Combined with a Lanczos algorithm, the BSE can be solved with a scaling of $\mathcal{O}(N_o N_u N_k \log N_k)$ (Henneke et al., 2020).

Statement of need

Due to the unfavorable scaling of solving the BSE directly, many interesting problems such as complex materials with large unit cells or systems requiring a dense Brillouin-zone sampling are not feasible. Even though Henneke and coworkers (Henneke et al., 2020) have already described the new algorithm and demonstrated the scaling improvement, an easy-to-use and scalable implementation was still missing. We have implemented and fully integrated this approach in the existing BSE infrastructure of the all-electron, full-potential package exciting (Gulans et al., 2014). Users can now easily choose which algorithm they prefer to use and have the full suite of exciton analysis implemented in exciting at hand.

Results

For computing the ISDF, two new parameters, $n_\mathbf{r}$ and c_μ, are introduced. $n_\mathbf{r}$ is the real-space sampling density for $u_{i\mathbf{k}}(\mathbf{r})$ and is defined as

$$n_\mathbf{r} = \frac{N_\mathbf{r}}{\Omega} \,,$$

where $N_\mathbf{r}$ is the number of $\mathbf{r}$-points and Ω the unit cell volume. The sampling is chosen to be regular such that the distance between the sampling points in each lattice direction is as similar as possible. The dimensionless parameter c_μ is used to control the number of interpolation points and is defined as

$$N_\mu = c_\mu \sqrt{\sqrt{N_\text{pairs}}} \,,$$

where N_pairs refers to the number of wave function pairs for which ISDF is computed. Note that N_pairs depends on $N_\mathbf{k}$. The double square-root dependence ensures that the overall scaling remains below $\mathcal{O}(N_\mathbf{k}^2)$. In Figure 1 we present, for the example of diamond, the difference in exciton binding energies obtained with the new implementation and a reference calculation. The reference, based on the direct implementation, sets up and diagonalizes the full BSH and depends neither on $n_\mathbf{r}$ nor c_μ. The results are shown as functions of $n_\mathbf{r}$ and c_μ for a small $\mathbf{k}$-grid of $2 \times 2 \times 2$. Additionally, we show the spectral similarities compared to the reference calculation as functions of $n_\mathbf{r}$ and c_μ. For both parameters, both properties converge as their values increase. To find the optimal interpolation grid for ISDF, we first converge $n_\mathbf{r}$, then

c_μ. In our example, $n_\mathrm{r} = 138$ [a.u.] and $c_\mu = 40.0$ yield converged results. This corresponds to a real-space sampling of $22 \times 22 \times 22$ and numbers of interpolation points $N_\mu^V = 202$, $N_\mu^{W_o} = 322$, and $N_\mu^{W_u} = 360$.

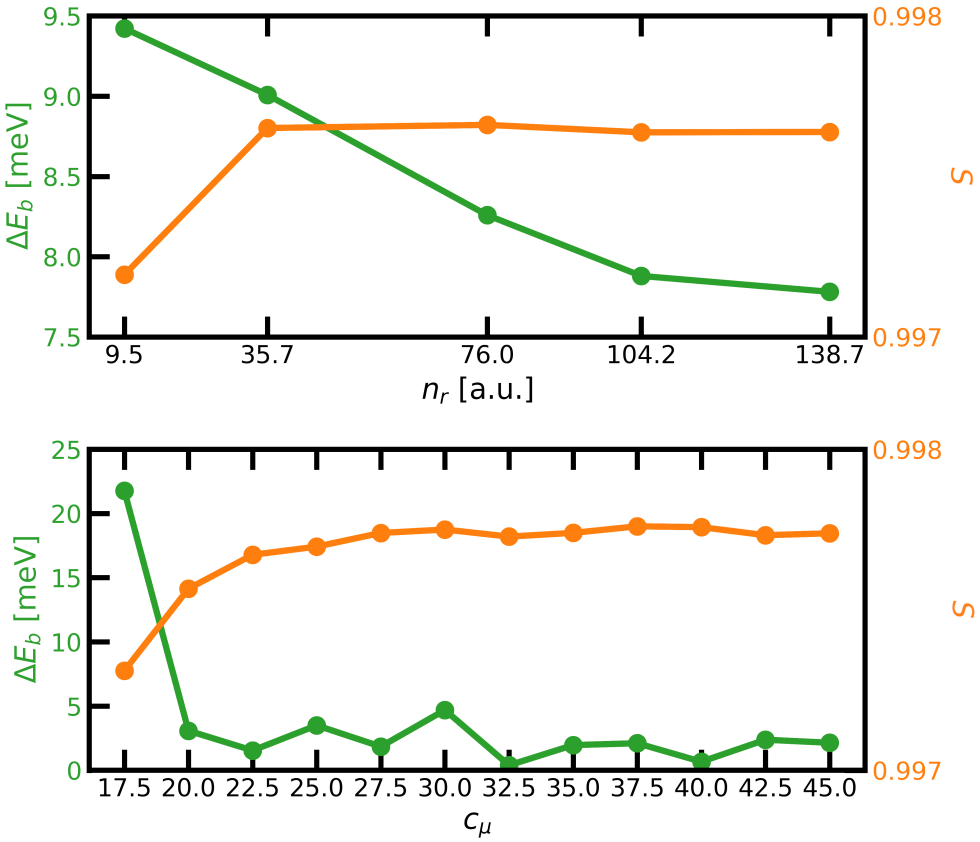

Figure 1: Difference of the exciton binding energy ΔE_b and spectral similarity S between the new method and the direct method as functions of n_r (upper panel) and c_μ (lower panel) for diamond on a $2 \times 2 \times 2$ $\mathbf{k}$-grid.

In Figure 2, we compare the exciton binding energies ΔE_b and spectra of the new implementation to those of the old implementation for increasing $\mathbf{k}$-grids, while keeping n_r as well as N_μ^V, $N_\mu^{W_o}$, and $N_\mu^{W_u}$ fixed at the values above. We observe that the results converge as $N_\mathbf{k}$ increases. Thus, the number of interpolation points is asymptotically independent of $N_\mathbf{k}$. A similar behavior of ISDF was observed in (Lu & Ying, 2016).

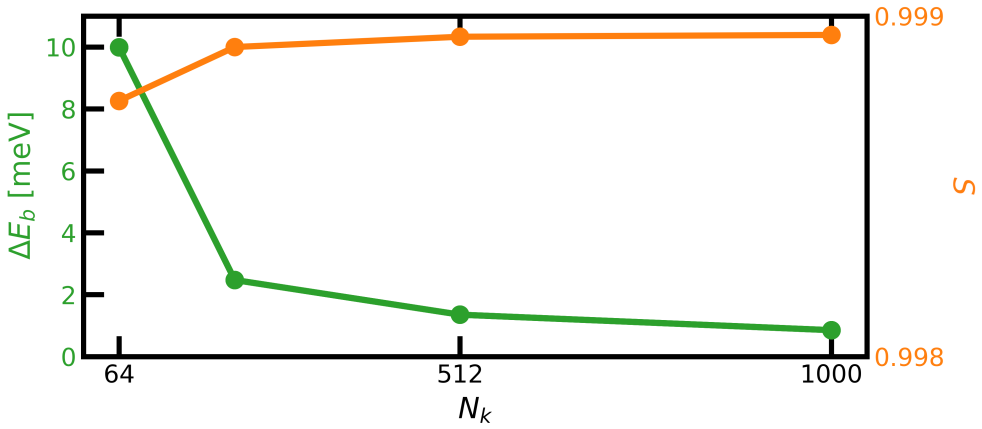

Figure 2: Difference of the exciton binding energy and spectral similarity between the new method and the direct method as functions of N_k for diamond.

In Figure 3, we show the wall times for solving the BSE with the new and direct implementations for increasing N_k. The new algorithm massively outperforms the direct one, and the speedup increases more than linearly with N_k. We also show the wall times for computing the RPA screening with increasing N_k, which is now clearly the bottleneck in solving the BSE.

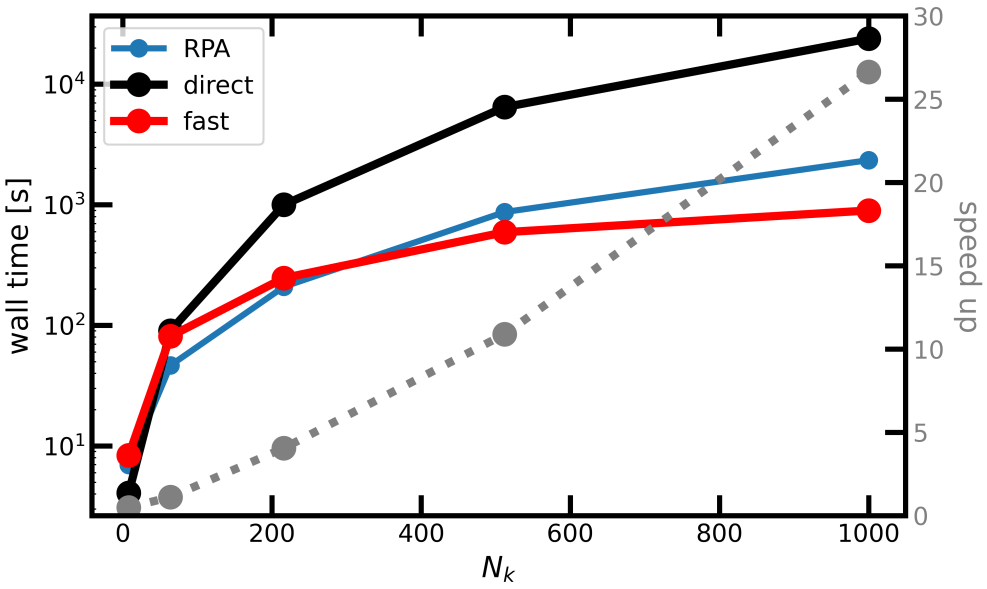

Figure 3: Runtimes of the direct (black) and new (red) BSE implementations and the RPA screening (blue) as functions of N_k. The speedup of the algorithm is shown by the gray dashed line.

Altogether, we have implemented a new, low-scaling BSE solver and fully integrated it into the all-electron, full-potential package exciting. We demonstrate that the new implementation yields results equivalent to the direct solution of the BSE but with significantly reduced computational time. Consequently, it enables more precise calculations and facilitates the study of more complex problems.

References

Dong, K., Hu, W., & Lin, L. (2018). Interpolative separable density fitting through centroidal Voronoi tessellation with applications to hybrid functional electronic structure calculations. *Journal of Chemical Theory and Computation*, *14*, 1311–1320. https://doi.org/10.1021/acs.jctc.7b01113

Gulans, A., Kontur, S., Meisenbichler, C., Nabok, D., Pavone, P., Rigamonti, S., Sagmeister, S., Werner, U., & Draxl, C. (2014). Exciting: A full-potential all-electron package implementing density-functional theory and many-body perturbation theory. *Journal of Physics: Condensed Matter*, *26*, 363202. https://doi.org/10.1088/0953-8984/26/36/363202

Henneke, F., Lin, L., Vorwerk, C., Draxl, C., Klein, R., & Yang, C. (2020). Fast optical absorption spectra calculations for periodic solid state systems. *Communications in Applied Mathematics and Computational Science*, *15*, 89–113. https://doi.org/10.2140/camcos.2020.15.89

Hu, W., Lin, L., & Yang, C. (2017). Interpolative separable density fitting decomposition for accelerating hybrid density functional calculations with applications to defects in silicon. *Journal of Chemical Theory and Computation*, *13*, 5420–5431. https://doi.org/10.1021/acs.jctc.7b00807

Lu, J., & Ying, L. (2015). Compression of the electron repulsion integral tensor in tensor hypercontraction format with cubic scaling cost. *Journal of Computational Physics*, *302*, 329–335. https://doi.org/10.1016/j.jcp.2015.09.014

Lu, J., & Ying, L. (2016). Fast algorithm for periodic density fitting for Bloch waves. *Annals of Mathematical Sciences and Applications*, *1*(2), 321–339. https://doi.org/10.4310/AMSA.2016.v1.n2.a3

Vorwerk, C., Aurich, B., Cocchi, C., & Draxl, C. (2019). Bethe–Salpeter equation for absorption and scattering spectroscopy: Implementation in the exciting code. *Electronic Structure*, *1*, 037001–037001. https://doi.org/10.1088/2516-1075/ab3123