pyGWBSE: a high throughput workflow package for GW-BSE calculations

Tathagata Biswas1 and Arunima K. Singh1,

We develop an open-source python workflow package, pyGWBSE to perform automated first-principles calculations within the GW-BSE (Bethe-Salpeter) framework. GW-BSE is a many body perturbation theory based approach to explore the quasiparticle (QP) and excitonic properties of materials. GW approximation accurately predicts bandgaps of materials by overcoming the bandgap underestimation issue of the more widely used density functional theory (DFT). BSE formalism produces absorption spectra directly comparable with experimental observations. pyGWBSE package achieves complete automation of the entire multi-step GW-BSE computation, including the convergence tests of several parameters that are crucial for the accuracy of these calculations. pyGWBSE is integrated with Wannier90, to generate QP bandstructures, interpolated using the maximally-localized wannier functions. pyGWBSE also enables the automated creation of databases of metadata and data, including QP and excitonic properties, which can be extremely useful for future material discovery studies in the field of ultra-wide bandgap semiconductors, electronics, photovoltaics, and photocatalysis.

ARTICLE OPEN

INTRODUCTION

Obtaining materials with properties that are optimized for a particular application traditionally relies on time-consuming and expensive experimental work. However, in recent years an alternative paradigm in the field of material discovery has emerged through the availability of modern massive supercomputing resources, development of first-principles methodologies, and ingenious computational algorithms. These advancements have pushed the boundaries of materials simulations, making them faster, more cost-effective, efficient and accurate. Consequently, high-throughput materials simulations have emerged as a tool for creating large databases and screening materials from these databases to identify candidate materials for applications in photocatalysis, energy storage, piezoelectrics, electrocatalysis etc.

However, applying similar approaches to applications related to optical and transport properties of materials have been hindered by a few technical obstacles. Density functional theory (DFT), the most widely used tool in computational high-throughput materials discovery studies are designed to explore ground state properties of a system and has been remarkably successful in predicting structural, mechanical, electronic, and thermal properties. However, studying the excited state properties of a system, such as optoelectronic or transport properties using DFT requires the interpretation of Kohn-Sham (KS) eigenvalues as energies involved in adding an electron to a many-electron system or subtracting one from it (QP energies). Following such a procedure, one often encounters the infamous bandgap underestimation problem due to the derivative discontinuity of the exchange-correlation energy.

A more rigorous approach to computing QP energies and accurate bandgap is applying many body perturbation theory (MBPT) within the GW approximation. Using this formalism, one computes the QP energies by calculating the first-order perturbative correction to the KS eigenvalues by approximating self-energy as a product of one particle Green’s function (G) and screened Coulomb interaction (W). It has been shown that MBPT within the GW approximation is particularly useful in computing QP properties of a wide variety of semiconductors and insulators without requiring any ad-hoc introduction of mixing parameters like those needed for hybrid functionals used in DFT.

Additionally, the study of QP properties using GW formalism enables us to compute several transport properties of materials that are inaccessible from a DFT calculation. For example, QP lifetimes calculated from GW calculations can be directly used to estimate impact ionization rates, a very useful parameter in the study of high-field transport of wide bandgap materials. In the case of low-field transport, in addition to the obvious importance of including GW corrections to KS eigenvalues and carrier mobilities, it has been shown that one needs to include the effects of GW correction on the orbital character of the relevant KS wavefunctions to obtain accurate transport properties of molecular junctions.

The optical and transport properties of semiconductors to a large extent are defined by the presence of intentional dopants or unintentional defects. GW formalism has emerged as a powerful approach that complements experiments and has become reliable enough to serve as a predictive tool for the crucial point defect properties such as charge transition levels and F-center photoluminescence spectra in semiconductors. Moreover, calculations based on MBPT using GW approximation has been successfully applied to estimate non-radiative recombinations such as Auger recombination rates which are very useful for optical applications. Auger recombination mechanism has been shown to cause significant efficiency loss in InGaN-based light-emitting diodes (LEDs), when operating at high injected carrier densities.

The necessity of the BSE methodology lies in the fact that even after including GW corrections the optical spectrum calculated within the independent-particle picture shows significant deviations from experimental results, as not only the absorption energies can be wrong, but often the oscillator strength of the...
peaks can deviate from the experiment by a factor of 2 or more. Moreover, it can not describe bound exciton states, which are particularly important in systems of reduced dimensions. The reason is, an independent particle picture can’t include electron-hole interactions (excitonic effects) which require an effective two-body approach. This can be achieved by evaluating the two-body Green’s function $G_2$ and formulating an equation of motion for $G_2$, known as the Bethe-Salpeter equation (BSE).

Despite its obvious indispensability, high-throughput computational material discovery studies for light-matter interaction related applications have rarely incorporated the QP or excitonic properties of materials using GW-BSE formalism mostly due to the unavailability of an automated workflow implementation that can perform such calculations. Two main challenges for such an implementation is the efficient convergence of multiple parameters and the tractability of the huge computational cost associated with the multi-step GW-BSE formalism. GW-BSE calculations are extremely sensitive to multiple interdependent convergence parameters such as the number of bands included in the GW self-energy calculation or the number of $k$-points used to sample the Brillouin zone (BZ) in the BSE calculation etc.

In this article, we introduce the open-source Python package, pyGWBSE, which automates the entire GW-BSE calculation using first-principles simulations software Vienna Ab-initio Software Package (VASP). This package enables automated input generation, submission to supercomputing platforms, analysis of post-simulation data, and storage of metadata and data in a MongoDB database. Moreover, pyGWBSE is capable of handling multiple convergence parameters associated with the GW-BSE formalism. Using this package, high-throughput computation of various electronic and optical properties is possible in a systematic and efficient manner. For example, the QP energies, bandstructures, and density of states can be computed using both the one-shot G0W0 and partially self-consistent GW0 level of GW formalism. The package enables automated BSE computations yielding the real and imaginary part of the dielectric function (incorporating electron-hole interaction), the exciton energies, and their corresponding oscillator strengths. DFT bandstructures, the orbital resolved density of states (DOS), electron/hole effective masses, band-edges, real and imaginary parts (absorption spectra) of the dielectric function ($\epsilon(\omega)$), and static dielectric tensors can also be computed using pyGWBSE.

The package is being continuously developed and the latest version can be obtained from the GitHub repository at https://github.com/cmdlab/pyGWBSE. pyGWBSE is built upon existing first-principles open source Python packages such as, pymatgen, Fireworks, and atomate. To obtain the QP bandstructure we use the Wannier90, a program for calculating maximally-localized Wannier functions to perform the interpolation required to obtain QP bandstructure with reduced computational cost.

pyGWBSE enables high-throughput simulations of highly reliable and efficient ab-initio approaches thus enabling future materials screening studies, the creation of large databases of high-quality computed properties of materials, and in turn machine learning model development. Hence, pyGWBSE could serve as a catapult to the next generation of technological advances in the field of power electronics, optoelectronics, photovoltaics, photocatalysis, etc.

In this article, we introduce the open-source Python package, pyGWBSE, which automates the entire GW-BSE calculation using first-principles simulations software Vienna Ab-initio Software Package (VASP). This package enables automated input generation, submission to supercomputing platforms, analysis of post-simulation data, and storage of metadata and data in a MongoDB database. Moreover, pyGWBSE is capable of handling multiple convergence parameters associated with the GW-BSE formalism. Using this package, high-throughput computation of various electronic and optical properties is possible in a systematic and efficient manner. For example, the QP energies, bandstructures, and density of states can be computed using both the one-shot G0W0 and partially self-consistent GW0 level of GW formalism. The package enables automated BSE computations yielding the real and imaginary part of the dielectric function (incorporating electron-hole interaction), the exciton energies, and their corresponding oscillator strengths. DFT bandstructures, the orbital resolved density of states (DOS), electron/hole effective masses, band-edges, real and imaginary parts (absorption spectra) of the dielectric function ($\epsilon(\omega)$), and static dielectric tensors can also be computed using pyGWBSE.

The package is being continuously developed and the latest version can be obtained from the GitHub repository at https://github.com/cmdlab/pyGWBSE. pyGWBSE is built upon existing first-principles open source Python packages such as, pymatgen, Fireworks, and atomate. To obtain the QP bandstructure we use the Wannier90, a program for calculating maximally-localized Wannier functions to perform the interpolation required to obtain QP bandstructure with reduced computational cost.

PyGWBSE enables high-throughput simulations of highly reliable and efficient ab-initio approaches thus enabling future materials screening studies, the creation of large databases of high-quality computed properties of materials, and in turn machine learning model development. Hence, pyGWBSE could serve as a catapult to the next generation of technological advances in the field of power electronics, optoelectronics, photovoltaics, photocatalysis, etc.

In the following sections, we present an overview of the underlying methodology, describe the workflow architecture, discuss the algorithms that were developed to perform the multi-step convergences, and benchmark the results obtained from the pyGWBSE workflow against experimental data in the literature.

### RESULTS AND DISCUSSION

**pyGWBSE: methodology and implementation**

The first-principles calculations in the pyGWBSE package are performed using one of the most well-known packages, VASP. VASP is a first principles computer program for atomic scale materials modeling. It is capable of using the projector-augmented wave (PAW) method and it comes with a rigorously tested pseudopotential library. VASP provides the accuracy of the full-potential linearized augmented plane wave (FLAPW) method but is computationally less expensive than most of the traditional plane wave-based methods. Currently, in addition to DFT, GW, and BSE, it supports various other post-DFT methods such as TD-DFT, ACFDT, 2nd-order Møller-Plesset perturbation theory and is under constant development, which opens up the possibility of implementing these methodologies in the pyGWBSE workflow in the future. Additionally, VASP is very efficiently parallelized and can utilize the potential of modern computers of both CPU and GPU-based architectures.

The pyGWBSE package is capable of computing several material properties using the DFT, GW, and BSE methodologies. Using pyGWBSE, properties such as bandstructures, the orbital resolved density of states (DOS), electron/hole effective masses, band-edges, real and imaginary parts (absorption spectra) of the dielectric function ($\epsilon(\omega)$), and static dielectric tensors can be computed using the DFT methodology. In the pyGWBSE package, the GW formalism can be used to compute QP energies both at one-shot G0W0 and partially self-consistent GW0 level of accuracy. The package uses maximally localized wannier functions (MLWF) to compute the electronic structure at the QP level of accuracy but at a significantly reduced computation cost. Using this package, the BSE methodology can be used to compute the real and imaginary parts of the dielectric function (incorporating electron-hole interaction), the exciton energies, and their corresponding oscillator strengths.

In the following sections, we describe the pyGWBSE package’s computational framework with a particular emphasis on all the crucial computational parameters which are needed to be converged for obtaining accurate results. Figure 1 provides a condensed diagrammatic representation of the GW-BSE framework described in the following two subsections; showing the interconnection between the key equations and the various physical quantities. Detailed discussions about the GW-BSE methodology can be found in several review articles. Additionally, we also present the rationale behind the strategies that we adopt to reduce the computational cost of the convergence calculations and thus make the computations more efficient.

**Fig. 1 A schematic summarizing the GW-BSE framework.** All the equations as well as the key quantities mentioned in GW and BSE methodology sections are drawn to provide a condensed visual representation.
GW methodology: converging QP properties

The QP energies are the energies for adding an electron to a many-electron system or subtracting one from it. Within the MBPT theory they can be calculated by solving the following equation:6

\[
(T + V_{n-e} + V_H - E_{\text{QP}}(\mathbf{r})) \psi_{\text{QP}}(\mathbf{r}) + \int d^3r' \Sigma(\mathbf{r}, \mathbf{r'}, E_{\text{QP}}) \psi_{\text{QP}}(\mathbf{r'}) = 0
\]

(1)

where, \(T\) is the kinetic energy operator, \(V_{n-e}\) is the operator to account for the nuclear(on)-electronic interaction, \(V_H\) is the Hartree potential, \(\Sigma\) is the self-energy operator and \(\mathbf{r}\) is the position vector of the electron. \(E_{\text{QP}}\) and \(\psi_{\text{QP}}\)'s are the QP energies and wavefunctions for \(n^{th}\) band with wavevector \(\mathbf{k}\).

GW approximation provides a practical route to compute the self-energy operator, \(\Sigma(\mathbf{r}, \mathbf{r'}, \omega)\), from KS wavefunctions, \(\psi_{\text{KS}}\) and energies, \(\epsilon_{\text{KS}}\) through one particle Green’s function, \(G(\mathbf{r}, \mathbf{r'}, \omega)\). Within the GW approximation \(\Sigma\) can be written as,

\[
\Sigma(\mathbf{r}, \mathbf{r'}, \omega) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} e^{i\mathbf{q}\cdot(\mathbf{r} - \mathbf{r'})} \delta(\omega) \frac{1}{|G\mathbf{q} + \mathbf{G}|} \left| \psi_{\text{KS}}(\mathbf{r}) \right|^2 \mathrm{d}\omega\]

(2)

where, \(W\) is the screened Coulomb interaction, \(\omega\) is the frequency and \(\delta\) is the positive infinitesimal.

\(W\) can be computed from \(\psi_{\text{KS}}\) and \(\epsilon_{\text{KS}}\) through independent particle polarization and frequency-dependent dielectric matrix utilizing the following three equations and taking advantage of the Random Phase Approximation (RPA).8

\[
W_q(\mathbf{G}, \mathbf{G}', \omega) = 4\pi e^2 \frac{1}{|\mathbf{q} + \mathbf{G}|} \epsilon_q^{-1}(\mathbf{G}, \mathbf{G}', \omega) \left| \mathbf{q} + \mathbf{G} \right|^2
\]

(3)

The dielectric matrix, \(\epsilon_q(\mathbf{G}, \mathbf{G}', \omega)\) is related to \(\chi\) as,8

\[
\epsilon_q(\mathbf{G}, \mathbf{G}', \omega) = \delta_{\mathbf{GG}'} - \frac{4\pi e^2}{|\mathbf{q} + \mathbf{G}|} \chi_q(\mathbf{G}, \mathbf{G}', \omega)
\]

(4)

Within RPA, independent particle polarization, \(\chi_q^0(\mathbf{G}, \mathbf{G}', \omega)\), is calculated as,8

\[
\chi_q^0(\mathbf{G}, \mathbf{G}', \omega) = \Omega^{-1} \sum_{\mathbf{n}, \mathbf{k}} 2w_\mathbf{k} (f_{\mathbf{n}, \mathbf{k}} - f_{\mathbf{n}, \mathbf{k}}) \times \frac{\langle \psi_{\mathbf{n}, \mathbf{k}} | e^{i\mathbf{q}\cdot\mathbf{r}} \psi_{\mathbf{n}, \mathbf{k}} \rangle}{\omega - \epsilon_q + \epsilon_q - \omega - i\eta} \delta(\omega_q - \epsilon_q)
\]

(5)

where, \(w_\mathbf{k}\) is the \(k\)-point weight, the \(f_{\mathbf{n}, \mathbf{k}}\) are the electron occupancy of the corresponding states, \(\mathbf{q}\) is the Bloch wave vector, \(\mathbf{G}\) is the reciprocal lattice vector and \(\eta\) is the infinitesimal complex shift.

Once the screened Coulomb interaction, \(W\), is computed, the diagonal matrix elements of self-energy operator, \(\Sigma(\omega)_{\mathbf{n}, \mathbf{k}, \mathbf{n}, \mathbf{k}}\) can be obtained using6

\[
\Sigma(\omega)_{\mathbf{n}, \mathbf{k}, \mathbf{n}, \mathbf{k}} = \Omega^{-1} \sum_{\mathbf{q}, \mathbf{G}} 2w_\mathbf{k} \int_0^{\infty} d\omega' W_q(\mathbf{G}, \mathbf{G}', \omega') \times \frac{\langle \psi_{\mathbf{n}, \mathbf{k}} | e^{i\mathbf{q}\cdot\mathbf{r}} \psi_{\mathbf{n}, \mathbf{k}} \rangle}{\omega - \epsilon_q - \omega'} \delta(\omega_q - \epsilon_q)
\]

(6)

where \(\mu\) is the Fermi energy. In the non-self-consistent GW calculation (also known as \(G_0W_0\) G0 and \(W_0\) are calculated using diagonal matrix elements of self-energy operator, \(\Sigma(\omega)_{\mathbf{n}, \mathbf{k}, \mathbf{n}, \mathbf{k}}\) approximated as the DFT wavefunction \(\psi_{\text{KS}}\).

Since Eq. (7) requires the values of \(E_{\text{QP}}\), the equation must be solved by iteration. Using the usual Newton-Raphson method for root finding, one can obtain the following update equation,8

\[
E_{\text{QP}}^{(n+1)} = E_{\text{QP}}^{(n)} + Z_{\text{nk}} \Re \left[ \langle \psi_{\text{nk}} | T + V_{n-e} + V_H + \Sigma(\mathbf{r}) | \psi_{\text{nk}} \rangle \right]
\]

(8)

where, \(Z_{\text{nk}}\) is the renormalization factor and can be calculated as,

\[
Z_{\text{nk}} = \left( 1 - \Re \left( \frac{\partial \Sigma(\omega) \psi_{\text{nk}}}{\partial \omega} \psi_{\text{nk}} \right) \right)^{-1}
\]

(9)

The iteration starts from DFT eigenvalues and if one stops after the first iteration the QP energies are obtained within \(G_0W_0\) approximation. One can also continue to obtain QP energies that are self-consistently converged. This scenario is referred to as self-consistent GW approximation (scGW).9

In the \(pyGW\) workflow, both the \(G_0W_0\) and scGW are implemented. In the partially self-consistent scGW approximation the \(\Sigma\) is updated self-consistently until convergence is reached but the \(W\) is kept unchanged, thus the scGW is also referred to as GW0. A full update of the \(\Sigma\) and \(W\) is seldom adopted and thus is not included in the \(pyGW\) package. In fact, it has been shown by several studies40-42 pertaining to free-electron gas, metals, and semiconductors that fully self-consistent GW calculations, without vertex corrections, lead to an overestimation of bandgaps.

There are three crucial parameters in a GW calculation that need to be converged to obtain accurate results, namely:

- Number of plane waves used to expand the screened Coulomb operator, \(W_q(\mathbf{G}, \mathbf{G}', \omega)\). This parameter can be specified by using ENCUTGW in the VASP implementation.
- Number of frequency grid points used in Eq. (6) for the frequency integration. This parameter can be specified by using NOLMGA in the VASP implementation.
- Number of bands used in Eqs. (5) and (6) for the summation. This parameter can be specified by using NBNDS in the VASP implementation.

In principle, the QP energies are to be converged w.r.t all three of the parameters mentioned above. Therefore, in our workflow, we have implemented the convergence tests for QP energies w.r.t these parameters namely, ENCUTGW, NOLMGA, and NBNDS.

Table 1 and 2 show the QP gaps of 9 materials computed with different values of ENCUTGW and NOLMGA. Table 1 shows the QP gaps computed with ENCUTGW values of 100, 150, and 200 eV, whereas Table 2 shows the QP gaps computed with NOLMGA values of 50, 65, and 80. From Table 1 we can see that, an ENCUTGW value of 150 eV is sufficient for obtaining a QP gap value that is converged within \(-0.1 eV\) for almost all the materials except diamond (C), for

Table 1. QP gaps of 9 materials computed using a screened coulomb cutoff, ENCUTGW, of 100 eV (\(E_{\text{100}}\)), 150 eV (\(E_{\text{150}}\)) and 200 eV (\(E_{\text{200}}\)).

| mp-id | Formula | \(E_{\text{100}}\) | \(E_{\text{150}}\) | \(E_{\text{200}}\) |
|-------|---------|----------------|----------------|----------------|
| mp-149 | Si      | 1.18           | 1.20           | 1.20           |
| mp-390 | TiO₂    | 3.89           | 3.89           | 3.89           |
| mp-66  | C       | 5.24           | 5.40           | 5.46           |
| mp-2133 | ZnO    | 2.68           | 2.68           | 2.67           |
| mp-804 | GaN     | 2.84           | 2.87           | 2.86           |
| mp-2624 | AlSb   | 1.69           | 1.69           | 1.70           |
| mp-1434 | MoS₂   | 2.33           | 2.34           | 2.34           |
| mp-984 | BN      | 5.45           | 5.46           | 5.46           |
| mp-22862 | NaCl  | 7.71           | 7.72           | 7.73           |

The table shows that a 200 eV cutoff is sufficient to converge QP gaps within \(0.1\ eV\) for all these materials. We have also listed the Materials Project identifier (mp-id) of all these materials, which can be used to obtain structural and DFT computed electronic properties of these materials.
which we need an ENCUTGW value of 200 eV. Table 2 shows that we need a NOMEQA of 80 to converge the QP gap within 0.1 eV for all the selected materials. Thus, as recommended by the VASP manual, ENCUTGW value of 2/3 × ENCUT and NOMEQA value of 50–100, would be sufficient to obtain accurate QP energies for a wide range of materials. Furthermore, based on the convergence of the 9 materials in Tables 1 and 2 we surmise that given the variety in the chemical compositions and crystal structures of these materials, it is likely that a value of 200 eV for ENCUTGW and 80 for NOMEQA may be sufficient to converge QP gaps of a variety of other materials within 0.1 eV.

The third parameter, NBANDS, however, needs to be converged for every material. Despite many efforts to eliminate or reduce the need of including a large number of empty orbitals in the GW calculation, it still is one of the major computation costs of a GW calculation. While several methods have been proposed to reduce the total number of empty orbitals in a GW calculation such as replacing actual KS orbitals with approximate orbitals generated using a reduced basis set, truncation of the sum over empty orbitals to a reduced number, and adding the contribution of the remaining orbitals within the static (COHSEX) approximation and modified static remainder approach. These methods are not currently implemented in the VASP package. However, in the future, if these methods are implemented, they can be easily incorporated into the pyGWHSB package and would reduce the computational costs of GW-BSE calculations significantly. Meanwhile, we strongly suggest performing convergence tests w.r.t NBANDS using pyGWHSB to obtain accurate results. It has been observed that convergence of all the aforementioned three parameters can be strongly interdependent and hence should be considered together.

However, the computational cost of treating multiple convergence parameters together scales quite rapidly. To address this issue we suggest that one should check the convergence of NBANDS first and once convergence is reached, increase ENCUTGW and NOMEQA by ~25% and check the convergence of the quasiparticle gap once again. For concurrent convergence of all the three parameters, the convparam tag in the input file can be specified as a list of [Δ NBANDS, Δ ENCUTGW, Δ NOMEQA], where Δ NBANDS, Δ ENCUTGW, and Δ NOMEQA are the increase in respective parameters in each convergence iteration.

### BSE methodology: converging excitonic properties

Studying optical electron-hole excitations is an effective two-body problem. In most cases single-particle picture of individual quasi-electron and quasi-hole excitations are not enough. We need to include electron-hole interactions as well. We can work with two-body Green’s function $G_2$, on the basis of the one-body Green’s function $G_1$, which can be described by the GW approximation. We can use QP electron and hole states of $G_1$, and their QP energies to estimate the electron-hole interactions. The equation of motion for $G_2$ is known as the Bethe-Salpeter equation and is very useful in the study of correlated electron-hole excitation states also known as excitons.

Following Strinati, Rohlffing and Louie, the BSE can be written as a generalized eigenvalue problem and the electron-hole excitation states can be calculated through the solution of BSE. For each exciton state $S$, within the Tamm-Dancoff approximation the BSE can be written as,

$$\langle \epsilon^{\text{QP}}_c - \epsilon^{\text{QP}}_d \rangle A_{vck} + \Sigma_{vck} (vck | \epsilon^{\text{QP}}_c | vck) = \Omega^2 A_{vck}$$

where, $A_{vck}$ is the exciton wavefunction, $\Omega^2$ is the excitation energy, $\epsilon^{\text{QP}}_c$ and $\epsilon^{\text{QP}}_d$ are the QP energies of the conduction ($vck$) and valence states ($|vck\rangle$) which is computed using the GW methodology discussed in the previous section. The electron-hole interaction kernel, $k^{\text{eh}}$, can be separated in two terms, $k^{\text{eh}} = K^d + K^r$, where $K^d$ is the screened direct interaction term and $K^r$ is the bare exchange interaction term. Within the GW approximation for $\Sigma$, in the basis of the single-particle orbitals in real space $(\phi_v | \phi_v) (k)$, the KS orbitals obtained from DFT calculations, they are defined in the following way,

$$\langle v \mid [K^d \mid v'] \rangle = \int d^3x d^3x^' \sum_{j} \delta(\mathbf{r} - \mathbf{r}') \phi_v^*(\mathbf{r}) \phi_v(\mathbf{r}) = \int d^3x d^3x^' \sum_{j} \delta(\mathbf{r} - \mathbf{r}') \phi_v(\mathbf{r}) \phi_v^*(\mathbf{r})$$

Once we have the solutions of the BSE Hamiltonian, we can construct $\epsilon_2(\omega)$ which incorporates excitonic effects from the solutions of the modified BSE,

$$\epsilon_2(\omega) = \sum_{k} \int d^3r \cdot |\psi(\mathbf{r})|^2 \delta(\omega - \Omega^2)$$

where $\Omega$ is the polarization vector, and $\psi$ is the velocity operator along the direction of the polarization of light, $\lambda$. The real part of the dielectric function $\epsilon_1(\omega)$, can be obtained by integration of $\epsilon_2(\omega)$ over all frequencies via Kramers-Kronig relations.

Converging the BSE absorption spectra with the number of $k$-points used to sample the BZ is one of the most computationally demanding tasks in a GW-BSE calculation. In this study, we propose a strategy to achieve this convergence with a significant reduction in the computational cost involved. We propose to obtain a convergent $\epsilon_2(\omega)$ within an independent particle picture (RPA) and use the same $k$-mesh to perform the BSE calculation. This strategy is expected to be useful because of the following reasons:

- There are two aspects of the convergence of $\epsilon_2(\omega)$ w.r.t. $\rho_k$. Firstly, due to band dispersion, one needs to include all the occupied-unoccupied transitions throughout the BZ to obtain absorption spectra that are converged. As a result, one needs to use a very dense $k$-grid for materials that have a stronger dispersion of bands near the gap. Secondly, the electron-hole interaction kernel is also dependent on the $k$-grid density, as the integrations described in Eqs. (11) and (12) are evaluated by a plane-wave summation in reciprocal space with the help of Fourier transform. Typically, the band dispersion doesn’t change significantly from DFT to GW bandstructure. Thus one should be able to estimate the $k$-grid density required to converge a BSE absorption spectra only by observing the change in the RPA
absorption spectra. For the convergence of electron-hole interaction kernel, we note that in the literature it has been suggested that $K^{\text{eh}}$ varies little w.r.t. the $k$-points, as the single-particle wave functions are quite robust w.r.t. $k$.

Therefore, one can assume that the $k$-mesh required to achieve the convergence solely from a change in band dispersion (which can be estimated from RPA calculations) is likely to be sufficient to converge $K^{\text{eh}}$ and also the BSE absorption spectra. In the following paragraph, we will discuss the results from our calculation, which supports the aforementioned hypothesis. Moreover, we want to emphasize that by convergent absorption spectra we mean that not only the positions of the absorption peaks are converged but also the peak positions and their oscillator strengths (peak heights) in $\varepsilon(\omega)$. In the case of Si, Diamond, and AlSb the consequences of including excitonic effects are relatively low as the relative heights of the low energy absorption peaks don’t change significantly (Fig. 2a–c and e–g), whereas, in the case of NaCl it shows a very prominent low energy excitonic peak, almost absent in the RPA spectra (Fig. 2d and h).

Figure 3a shows the SC for the spectra of Fig. 2 computed with a $\Delta\rho_k$ of 50 Å$^3$. The SC captures both the shifts in peak positions and oscillator strengths in the absorption spectra resulting from the change in reciprocal density used in BSE calculations. As one can see from the absorption spectra of Si and AlSb, Fig. 2a and c, that the peak positions and their oscillator strengths don’t change too much when $\rho_k > 100$ Å$^3$. This is reflected by a larger value of SC > 0.75, for them (Fig. 3a). Whereas, a lower SC value is obtained for diamond and especially for NaCl even for $\rho_k > 100$ Å$^3$. This is mostly due to a large change in peak positions and a change in oscillator strengths for NaCl (Fig. 2d and h). Nevertheless, once we look at the convergence of SC for both BSE and RPA absorption spectra (Fig. 3a), they look quite similar. Thus the lower resource and time-intensive RPA can be employed to estimate the $k$-mesh required to converge the BSE calculations (see Supplementary Figure 1 for the computational cost comparison). In Fig. 3b we show the convergence of optical gap, $E_{\text{opt}}$, with the reciprocal density, $\rho_k$, used in BSE calculation. To compare the convergence for these materials that have very different optical gaps we subtracted the converged value of the optical gap, $E_{\text{opt}}$, for each of them to show the variation in the same scale. From Fig. 3b it is clear that the $k$-mesh required to obtain an SC > 0.75 is significantly. In that case, one should not use the RPA
This page discusses the calculation and analysis of the Bethe-Salpeter equation (BSE) for semiconductor materials, focusing on the convergence of BSE calculations and the use of a workflow architecture to automate this process. The BSE is used to obtain quasiparticle (QP) energies along the high-symmetry k-path. The workflow architecture is implemented using the Fireworks software, which includes seven Fireworks (FWs) designed to perform different tasks such as structural relaxation, effective mass calculations, and wannier interpolation. The FWs are executed in a specific order to achieve the desired convergence. The results are stored in a MongoDB database for future reference and analysis. The workflow is designed to be flexible and scalable, allowing for the addition of new FWs as needed. The example 'input.yaml' file is used to define the input parameters for each FW, and the workflow architecture is demonstrated using graphs and tables.
software. The third category of FTs, communication FTs (shown in light purple boxes in Fig. S3), enables the communication between different FTs. For example, PasscalcLocs in FW1 and FW5 is used to pass the address of the directory where a parent FW was executed to its children’s FWs. The last category of FTs, the transfer to database FTs, transfers information to the database. They are shown by gray boxes. For example, Eps2db in FW1 is used to read the dielectric tensor from VASP output file and save it to the database.

We use separate MongoDB collections to store the data and metadata, including inputs and outputs, associated with the FWs and FTs. Figure 4 shows these various collections. The group of all these collections is called the GWBSE_DB. Dielectric tensors, KS eigenvalues, projections of the KS wavefunctions onto atomic orbitals (for computing projected DOS), RPA dielectric functions from DFT calculation, effective masses, wannier interpolated bandstructures from both DFT and GW levels, all the QP energies including those during convergence, frequency-dependent dielectric function for different light polarization axis from BSE calculations are some of the key quantities stored in the collections.

A more elaborate description of the code’s features and the functionality implementation, especially the convergence of the GW and BSE-related parameters can be found in the Supplementary Discussion and Supplementary Notes section of the Supplementary Information. Moreover, the Supplementary Information (Supplementary Methods) includes an example Jupyter Notebook that shows a step-by-step setup process to create a workflow and analyze the results obtained from the workflow to determine QP properties and the BSE absorption spectra of the wurtzite phase of AlN.

**Benchmarking pyGWBSE for wurtzite AlN**

Recently, we employed pyGWBSE to compute the excitonic effects in absorption spectra of ~50 photocatalysts using the Bethe-Salpeter formalism. In that study, we have compared the QP gap computed using pyGWBSE for 10 materials of very different chemical compositions with the experimental values for the purpose of benchmarking and found good agreement. However, in the aforementioned study, we haven’t utilized all the functionalities of the pyGWBSE workflow by applying it to a test case of wurtzite-AlN. In this section, we compare the various quantities obtained from the workflow simulations with the experimentally measured values reported in the literature.

We begin by evaluating the quantities obtained from the DFT calculations. To that end, the dielectric constants of AlN are found to be \( \varepsilon_{\text{KS}} = 4.61 \), \( \varepsilon_{\text{KS}} = 4.82 \), and \( \varepsilon_{\text{KS}} = 4.68 \), with the \( \varepsilon_{\text{KS}} \) being in exact agreement with the experimentally measured value of 4.68.

The electron effective masses are found to be \( m_{e} = 0.28m_{0} \) and \( m_{e} = 0.30m_{0} \) which fall in the experimentally obtained ranges of ~ 0.29–0.45\( m_{0} \). In case of hole effective masses we find a large anisotropy, \( m_{h} = 0.24m_{0} \) and \( m_{h} = 4.32m_{0} \). The average hole effective mass in AlN was recently estimated to be ~2.7\( m_{0} \), based on the available experimental data.
on experimental measurements of the Mg acceptor binding energy in Mg-doped AlN epilayers \(^{33}\) which is very similar to our average computed value of \(m_{\text{avg}}^2 = 2.96\).

Figure 5a shows the bandstructure of AlN computed from DFT using VASP directly (red solid) and through the use of wannier interpolation (black dashed). As one can see that, the wannier interpolation is very accurate and both the bandstructures overlap with each other throughout most of the BZ. Figure 5a also shows that AlN is a direct gap semiconductor with a DFT gap of 4.05 eV, which is expectedly underestimated compared to the experimental value of 6.2 eV \(^{54}\) but very close to the value obtained from the literature \(^{56}\). The light polarization is set to be along the [0001] direction which enables high-throughput GW-BSE calculations. In this article, we present the underlying theory, the work function, the algorithmic implementation, and benchmark simulations to even better agreement with the experimental spectra \(^{57}\).

Once we perform the one-shot GW calculation the direct gap increases to 5.59 eV. Although this is closer to the experimental value it is still not quite accurate. Previous studies, using an LDA functional as a starting point found a QP gap of 5.8 eV, which also doesn’t agree with the experimental gap. However, after we perform partial self-consistent GW (scGW) the QP gap becomes 6.28 eV, resulting in an good agreement with the experimental value of 6.2 eV \(^{54}\). The QP bandstructure with scGW is shown by the blue curve in Fig. 5a.

We use the QP energies and projection of KS wavefunctions or atomic orbitals to compute the orbital resolved DOS with QP corrections under the assumption that the KS wavefunction is a good approximation for the QP wavefunction. Figure 5b shows the orbital-resolved DOS of wurtzite-AlN with Al(p) and N(p) states shown with dashed and dotted lines respectively. We show orbital-resolved DOS obtained from both DFT and GW calculations for comparison. Our calculation suggests that the valence band edge of wurtzite-AlN mostly consists of N(p) states whereas the conduction band edge is resulting from strong hybridization between Al(p) and N(p) states, which is consistent with the findings of previous studies \(^{55}\).

To show pyGWBSE’s ability to perform BSE calculation and obtain absorption spectra \(\epsilon_{2}(\omega)\) that include electron-hole interactions we perform the GW-BSE calculation for wurtzite-AlN. Figure 5c compares the absorption spectra \(\epsilon_{2}(\omega)\) that we obtained from the BSE calculation, the calculation without electron-hole interaction, and the experimentally obtained spectra from the literature \(^{66}\). The light polarization is set to be perpendicular to the c-axis. As we can see from Fig. 5c the absorption spectra calculated without taking electron-hole (e-h) interaction into account completely misses the features in the 6–10 eV range, visible in the experimental absorption spectra (shown with ‘+’ symbols in Fig. 5c). Only when we include the e-h interaction through the BSE calculation those excitonic features are retrieved. Although, the absorption spectra obtained from BSE very closely resemble the experimental absorption spectra we find that the sharp absorption edge at 6.2 eV and two prominent absorption peaks at 7.85 and 8.95 eV are shifted (by ~ 0.15 eV for the peaks) to the lower frequencies. In the Supplementary Fig. 5 and Supplementary Table 1 in the SI, we can see that with an SC of ~ 0.8 reasonable amount of information is obtained for the spectra peaks and their positions, however, larger SCs lead to better accuracy, albeit at a higher computational cost and computing time. Thus previous GW-BSE calculations with a finer sampling of the BZ, with randomly distributed 1000 k-points, led to even better agreement with the experimental spectra \(^{57}\).

To conclude, we have developed a Python toolkit, pyGWBSE, which enables high-throughput GW-BSE calculations. In this article, we present the underlying theory, the workflow architecture, the algorithmic implementation, and benchmark simulations for the pyGWBSE code. This open-source code (available at https://github.com/cmdlab/pyGWBSE) enables automated input file generation, submission to supercomputing platforms, analysis of post-simulation data, and storage of metadata and data in a MongoDB database. Moreover, pyGWBSE is capable of handling multiple convergence parameters associated with the GW-BSE formalism. To reduce the computational cost associated with obtaining a converged absorption spectrum from BSE calculations, we present a strategy for computing the similarity coefficient from RPA spectra. We have shown that this approach ensures convergence of not only the optical gap or exciton binding energy but the entire absorption spectra in the desired frequency range. Our openly available code will help to include QP properties and excitonic effects in future computational material design and discovery studies in a variety of fields such as power electronics, photovoltaics, and photocatalysis. The pyGWBSE will facilitate high-throughput GW-BSE simulations enabling the application of large data methods to further explore our understanding of materials as well as first-principles methods that are designed for computing excited state properties.
METHODS
A detailed discussion of the convergence parameters for the DFT ground state calculations can be found in the existing literature\textsuperscript{18,19}. In these studies, plane wave energy cutoff was set to 1.3 times the maximum energy cutoff specified in the pseudopotentials and k-grid set as 500 n\textsuperscript{-1} points, where n represents the number of atoms in the unit cell distributed as uniformly as possible in k-space (see Supplementary Fig. 2 for convergence results). It resulted in total energy convergence of 15 meV atom\textsuperscript{-1} for 96% of 182 chemically diverse materials\textsuperscript{18}. The choice of exchange-correlation functional used in the pyGWBSE workflow is determined by the settings of VASP via the pymatgen\textsuperscript{21} installation and can be changed by following installation instructions of pymatgen package. Our convergence calculations performed on several materials (Supplementary Fig. 2) shows that, the same choice for the plane wave energy cutoff but a higher k-grid density is required to converge the GW-BSE calculations.

A broadening of 0.1 eV is used for calculating all the BSE as well as RPA absorption spectra reported in Fig. 2. Moreover, the number of valence (v) and conduction (c) bands included in the BSE calculation are selected to ensure all the vertical v→c transitions of energy <3 eV are included. One can change this energy window in the pyGWBSE workflow by using the enwinbs tag in the input file. In the benchmarking calculations for w-AlN we have used a reciprocal density, \( \rho_{R} \) value of 200 Å\textsuperscript{3} (12 × 12 × 7 k-grid) with a broadening (\textsc{cshift})\textsuperscript{28} of 0.2 eV, which produces an SC value of 0.91 (see Supplementary Fig. 5 and Supplementary Table 1 in Supplementary Information for detailed convergence results and associated computational cost requirements).

DATA AVAILABILITY
All data supporting the findings of this work are available in the paper and its Supplementary Information. Extra data and machine-readable data are available upon reasonable request to the authors.

CODE AVAILABILITY
The pyGWBSE code is being continuously developed and the latest version can be obtained from the GitHub repository at https://github.com/cmdlab/pyGWBSE.

Received: 11 October 2022; Accepted: 25 January 2023; Published online: 13 February 2023

REFERENCES
1. Singh, A. K., Montoya, J. H., Gregoire, J. M. & Persson, K. A. Robust and synthetizable photocatalysts for CO\textsubscript{2} reduction: a data-driven materials discovery. Nat. Commun. 10, 1–9 (2019).
2. Wu, Y., Lai, P., Haueter, G., Persson, K. & Ceder, G. First principles high-throughput screening of oxynitrides for water-splitting photocatalysts. Energy Environ. Sci. 6, 157–168 (2013).
3. Kirklin, S., Meredith, B. & Wolferton, C. High-throughput computational screening of new Li-ion battery anode materials. Adv. Energy Mater. 3, 252–262 (2013).
4. Haueter, G. et al. Phosphates as lithium-ion battery cathodes: an evaluation based on high-throughput ab initio calculations. Chem. Mater. 23, 3495–3508 (2011).
5. Choudhary, K. et al. High-throughput density functional perturbation theory and machine learning predictions of infrared, piezoelectric, and dielectric responses. npj Comput. Mater. 6, 19 (2020).
6. Greeley, J., Jaramillo, T. F., Borde, J., Chorkendorff, I. & Narowski, J. K. Computational high-throughput screening of electrocatalytic materials for hydrogen evolution. Nat. Mater. 5, 909–913 (2006).
7. Jones, R. O. Density functional theory: Its origins, rise to prominence, and future. Rev. Mod. Phys. 87, 897 (2015).
8. Perdew, J. P. Density functional theory and the band gap problem. Int. J. Quant. Chem. 28, 497–523 (1985).
9. Hedin, L. New method for calculating the one-particle Green’s function with application to the electron-gas problem. Phys. Rev. 139, 796 (1965).
10. Onida, G., Reining, L. & Rubio, A. Electronic excitations: density-functional versus many-body green’s-function approaches. Rev. Mod. Phys. 74, 601 (2002).
11. Muscat, J., Wander, A. & Harrison, N. On the prediction of band gaps from hybrid functional theory. Chem. Phys. Lett. 342, 397–401 (2001).
12. Vines, F., Lamiel-García, O., Chul Ko, Y., Yong Lee, J. & Illas, F. Systematic study of the effect of hse functional internal parameters on the electronic structure and band gap of a representative set of metal oxides. J. Comput. Chem. 38, 781–789 (2017).
13. Kotani, T. & Van Schilfgaarde, M. Impact ionization rates for Si, GaAs, InAs, InS, and GaN in the GW approximation. Phys. Rev. B 81, 125201 (2010).
14. Darancet, P., Ferretti, A., Mayou, D. & Olevano, V. Ab initio GW electron-electron interaction effects in quantum transport. Phys. Rev. B 75, 075102 (2007).
15. Rangel, T., Ferretti, A., Trevisanutto, P., Olevano, V. & Rignanese, G.-M. Transport properties of molecular junctions from many-body perturbation theory. Phys. Rev. B 84, 045426 (2011).
16. Biswas, T. & Jain, M. Electronic structure and optical properties of F-centers in aluminas. Phys. Rev. B 69, 144102 (2004).
17. Freysoldt, C. et al. First-principles calculations for point defects in solids. Rev. Mod. Phys. 86, 253 (2014).
18. Kioupakis, E., Rinke, P., Delaney, K. T. & Van de Walle, C. G. Indirect auger recombination as a cause of efficiency droop in nitride light-emitting diodes. Appl. Phys. Lett. 98, 161107 (2011).
19. McAllister, A., Åberg, D., Schleife, A. & Kioupakis, E. Auger recombination in sodium-iodide scintillators from first principles. Appl. Phys. Lett. 106, 141901 (2015).
20. Ugeda, M. M. et al. Giant bandgap renormalization and excitonic effects in a monolayer transition metal dichalcogenide semiconductor. Nat. Mater. 13, 1091–1095 (2014).
21. Rohlfing, M. & Louie, S. G. Electron-hole excitations and optical spectra from first-principles. Phys. Rev. B 62, 4927 (2000).
22. Hastrup, S. et al. The computational 2D materials database: high-throughput modeling and discovery of atomically thin crystals. 2D Mater. 5, 042002 (2018).
23. Rajan, A. C. et al. Machine-learning-assisted accurate band gap predictions of functionalized mxenes. Chem. Mater. 30, 4031–4038 (2018).
24. Liang, J. & Zhu, X. Phillips-inspired machine learning for band gap and exciton binding energy prediction. J. Phys. Chem. Lett. 10, 5640–5646 (2019).
25. Liu, X. et al. Finding predictive models for singlet fission by machine learning. npj Comput. Mater. 8, 1–10 (2022).
26. Hafner, J. Ab-initio simulations of materials using vasp: Density-functional theory and beyond. J. Comput. Chem. 29, 2044–2078 (2008).
27. Ong, S. P. et al. Python materials genomics (pymatgen): A robust, open-source python library for materials analysis. Comput. Mater. Sci. 68, 314–319 (2013).
28. Jain, A. et al. Fireworks: A dynamic workflow system designed for high-throughput applications. Concurr. Comput. Pract. Exp. 27, 5037–5059 (2015).
29. Mathew, K. et al. Atomeate: A high-level interface to generate, execute, and analyze computational materials science workflows. Comput. Mater. Sci. 139, 140–152 (2017).
30. Mostofi, A. A. et al. WANNIER90: A tool for obtaining maximally-localised wannier functions. Comput. Phys. Commun. 178, 685–699 (2008).
31. Bölöni, P. E. Projector augmented-wave method. Phys. Rev. B 50, 17953 (1994).
32. Singh, D. Planes Waves, Pseudopotentials and the LAPW Method. (Kluwer Academic, 1994).
33. Leng, X., Jin, F., Wei, M. & Ma, Y. GW method and Bethe–Salpeter electron equation for calculating electronic excitations. Wiley Interdiscip. Rev.: Comput. Mol. Sci. 6, 352–550 (2016).
34. Faber, C., Boulanger, P., Attaccalite, C., Duchemin, I. & Blase, X. Exotic states properties of organic molecules: From density functional theory to the GW and Bethe–Salpeter Green’s function formalisms. Philos. Trans. R. Soc. 372, 20130271 (2014).
35. Blase, X., Duchemin, I. & Jacquemin, D. The Bethe–Salpeter equation in chemistry: relations with TD-DFT, applications and challenges. Chem. Soc. Rev. 47, 1022–1043 (2018).
36. Hybertsen, M. S. & Louie, S. G. Electron correlation in semiconductors and insulators: Band gaps and quasiparticle energies. Phys. Rev. B 34, 5390 (1986).
37. Shishkin, M. & Kresse, G. Implementation and performance of the frequency-dependent GW method within the PAW framework. Phys. Rev. B 74, 035101 (2006).
38. Hybertsen, M. S. & Louie, S. G. Ab initio static dielectric matrices from the density-functional approach. I. formulation and application to semiconductors and insulators. Phys. Rev. B 35, 5585 (1987).
39. Shishkin, M. & Kresse, G. Self-consistent GW calculations for semiconductors and insulators. Phys. Rev. B 75, 235102 (2007).
40. Delaney, K., García-González, P., Rubio, A., Rinke, P. & Godby, R. W. Comment on “band-gap problem in semiconductors revisited: effects of core states and many-body self-consistency”. Phys. Rev. Lett. 93, 249701 (2004).
ACKNOWLEDGEMENTS

This work was supported by ULTRA, an Energy Frontier Research Center funded by the U.S. Department of Energy (DOE), Office of Science, Basic Energy Sciences (BES), under Award # DE-SC0021230. In addition, Singh acknowledges support by the Arizona State University start-up funds. The authors acknowledge the San Diego Supercomputer Center under the NSF-XSEDE Award No. DMR1500006 and the Research Computing at Arizona State University for providing HPC resources. This research used resources of 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. The authors also thank Tara M. Boland, Adway Gupta, Akash Patel, and Cody Milne for testing the code and for helpful discussions.

AUTHOR CONTRIBUTIONS

T.B. developed the pyGWBSE code and performed the simulations and calculations. Both authors contributed to designing the research methods, interpreting the data, and writing the paper.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41524-023-00976-y.

Correspondence and requests for materials should be addressed to Arunima K. Singh.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.