Static subspace approximation for the evaluation of G_0W_0 quasiparticle energies within a sum-over-bands approach

Mauro Del Ben,[1,*] Felipe H. da Jornada,[2,3] Gabriel Antonius,[2,3,4] Tonatiuh Rangel,[5,2]
Steven G. Louie,[2,3] Jack Deslippe,[6] and Andrew Canning[1]

[1]*Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*
[2]*Department of Physics, University of California at Berkeley, California 94720, USA*
[3]*Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*
[4]*Département de Chimie, Biochimie et Physique, Institut de recherche sur l'hydrogène, Université du Québec à Trois-Rivières, C.P. 500, Trois-Rivières, Canada*
[5]*Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*
[6]*NERSC, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*

(Received 4 December 2018; published 18 March 2019)

Many-body perturbation theory within the *GW* approach has been established as a quantitatively accurate approach for predicting the quasiparticle and excited-state properties of a wide variety of materials. However, the successful application of the method is often complicated by the computational complexity associated with the evaluation and inversion of the frequency-dependent dielectric matrix $\varepsilon(\omega)$. Here, we describe an approach to speed up the evaluation of the frequency-dependent part of $\varepsilon(\omega)$ in the traditional sum-over-states *GW* framework based on the low-rank approximation of the static dielectric matrix, a technique often used in *GW* implementations that are based on a starting mean field within density-functional perturbation theory. We show that the overall accuracy of the approach, independently from other calculation parameters, is solely determined by the threshold on the eigenvalues of the static dielectric matrix, $\varepsilon(\omega = 0)$, and that it can yield orders-of-magnitude speed-ups in full-frequency *GW* calculations. We validate our implementation with several benchmark calculations ranging from bulk materials to systems with reduced dimensionality, and show that this technique allows one not only to study larger systems, but also to carefully consider the convergence of computationally demanding systems, such as ZnO, without relying on plasmon-pole models.

DOI: 10.1103/PhysRevB.99.125128

I. INTRODUCTION

Throughout the last few decades, Kohn-Sham (KS) density-functional theory (DFT) has become a standard computational method for predicting properties of extended materials, nanostructures, and even complex molecules. With the recent improvements in energy functionals and numerical algorithms in DFT codes, it is now possible to accurately predict many ground-state properties of systems with complex chemical compositions and with thousand of atoms in the supercells.

Despite this notable success, static DFT is a ground-state formalism and does not yield accurate quasiparticle properties, such as quasiparticle energies, electronic gaps, effective masses, and quasiparticle lifetimes. A rigorous and proven approach for computing these properties is through many-body perturbation theory, in which the electronic Green's function G, the poles of which are associated with quasiparticle excitation energies, is written in terms of a noninteracting Green's function G_0 and the electronic self-energy Σ, the latter able to capture electron-electron interactions and eventually electron-phonon interactions. There are various prescriptions for approximating Σ; a particularly common and robust one that yields accurate quasiparticle properties for confined and extended systems—both metallic and semiconducting—is the *GW* approximation, in which the electronic self-energy is written as a product of the Green's function and the screened Coulomb interaction W [1,2], which we define in the next section.

The *ab initio GW* [2] approach is a particular way of using the *GW* approximation in first-principles calculations, where one typically uses KS orbitals and eigenvalues to construct G_0. The *ab initio GW* approach has been shown to yield excellent quasiparticle excitation properties for a variety of materials, and has received considerable attention from the computational sciences community. In fact, *GW* calculations have now been implemented in a number of software packages [3–13].

Despite the value of *GW* calculations, their usage is still less popular compared to DFT codes due to the significantly higher computational cost. The main bottleneck in traditional *GW* calculations, i.e., those which explicitly write the Green's function in terms of its spectral representation, is in the evaluation of the noninteractive polarizability matrix $\chi^0(\omega)$. The computational cost of computing $\chi^0(\omega)$ scales as $O(N^4)$, where N is proportional to the system size. In addition, each matrix has to be computed and eventually inverted on a number of distinct frequency points which typically ranges from a few tens up to hundreds.

One of the earliest strategies developed to reduce this computational bottleneck is to simplify the frequency dependence of $\chi^0(\omega)$, which allows one to perform a calculation of only the static polarizability $\chi^0(\omega = 0)$, and treat

*mdelben@lbl.gov

all integrals involving the frequency dependence of $\chi^0(\omega)$ analytically. Such approaches are referred to as generalized plasmon-pole models (PPMs), and were validated on a variety of systems [1,2,14–16].

Despite the success of PPMs, the simplification in the dynamical treatment of the polarizability typically prevents these approaches from accurately capturing dynamical effects of the self-energy. As a result, PPMs do not give accurate information about quasiparticle lifetimes. In addition, there is a growing demand for *GW* calculations with increasing accuracy, and PPMs may introduce uncontrollable approximations beyond the desired threshold, in particular on systems where the polarizability is not well approximated by a set of plasmon-like excitations. Thus, there is a renewed interest in *GW* calculations that do not employ PPMs, which are commonly referred to as full-frequency (FF) calculations.

A number of novel ideas have been proposed to alleviate the computational bottleneck in FF calculations. For instance, Liu and coworkers recently presented a cubic-scaling *GW* method based on a real-space, imaginary-time representation of the polarizability matrix and self-energy, which are then transformed to the real frequency domain through a nonuniform Fourier transform and analytical continuation [17]. In addition, linear-scaling stochastic *GW* approaches have also been demonstrated [18], which are particularly promising for systems with very large number of atoms. Even though approaches which reduce the scaling cost of *GW* calculations are an important avenue of research, there is still high value in speeding up traditional *GW* calculations evaluated directly on the real frequency axis, in particular if one is interested in resolving sharp features in the spectral function.

A more general way to speed up FF-*GW* calculations that is both elegant and powerful is to find a better basis set to describe the polarizability of the system [19,20]. This idea is motivated by the properties of the eigenspectrum of the dielectric function [21–23] and was shown to work remarkably well for *GW* calculations based on density-functional perturbation theory (DFPT) [4,24,25], as studied in detail in the earlier work of Wilson *et al.* [26,27], Nguyen *et al.* [28], and Pham *et al.* [29]. These studies have shown that, when solving the Sternheimer equation to obtain the polarizability matrix of a system, the eigendecomposition of $\chi^0(\omega = 0)$ forms a very good basis set for subsequent Sternheimer-equation calculations for $\omega \neq 0$. A natural question is if this technique can be used to speed up *GW* calculations that do not rely on DFPT.

In this paper, we describe the implementation of a method which greatly speeds up the evaluation of the frequency-dependent part of the polarizability, making the cost of a FF calculation of the same order as that based on a PPM. This is achieved through a low-rank approximation [30] of the static symmetrized susceptibility [28] defining a static basis set, and using this basis to perform a compression of the matrix elements involved in the computation of the polarizability matrix for the remaining frequencies. This procedure is efficiently executed on modern computer architectures through cache-friendly matrix-matrix multiplications and scales well to handle large systems [31].

We refer to this approach in the context of traditional *GW* calculations as the static subspace approximation, and we show that it performs remarkably well for all classes of systems tested in this work, with a large speedup, and with a small and controllable error which is determined by the eigenvalue threshold parameter. We also apply the method to study the quasiparticle band gap of zinc oxide, which is a computationally challenging system for *GW* calculations, partly because it is not well described with a plasmon-pole model. The new method has been implemented in the BERKELEYGW software package [3].

The paper is organized as follows. Section II reviews the key components required for practical *GW* calculations; Sec. III introduces our implementation of the static subspace approximation; Sec. IV shows validations of our method for a variety of systems and Sec. V reports on the application of the method for the evaluation of the quasiparticle band gap of ZnO.

II. SUM-OVER-BANDS GW APPROACH IN PRACTICE

In typical *ab initio GW* calculations, the noninteracting Green's function G_0 is constructed from the KS eigenvalues $\epsilon_{n\mathbf{k}}$ and wave functions $\psi_{n\mathbf{k}}(\mathbf{r})$, where n is a band index and $\mathbf{k}$ is a k point. This allows one to express the independent-particle irreducible polarizability χ^0 at zero temperature for an insulating system within the time-dependent Hartree, or random-phase approximation (RPA) [32–34], as

$$\chi^0(\mathbf{r},\mathbf{r}',\mathbf{q},\omega) = 2\sum_{n}^{N_v}\sum_{m}^{N_c}\int \frac{d\mathbf{k}}{(2\pi)^3}\,\Delta_{nm\mathbf{k}}(\mathbf{q},\omega) \times \psi_{n\mathbf{k}+\mathbf{q}}(\mathbf{r})\psi^*_{m\mathbf{k}}(\mathbf{r})\psi_{m\mathbf{k}}(\mathbf{r}')\psi^*_{n\mathbf{k}+\mathbf{q}}(\mathbf{r}'), \quad (1)$$

$$\Delta_{nm\mathbf{k}}(\mathbf{q},\omega) \equiv \frac{1}{\omega - (\epsilon_{m\mathbf{k}} - \epsilon_{n\mathbf{k}+\mathbf{q}}) + i\,\delta} - \frac{1}{\omega + (\epsilon_{m\mathbf{k}} - \epsilon_{n\mathbf{k}+\mathbf{q}}) - i\,\delta}, \quad (2)$$

where N_v and N_c denote the number of valence and conduction states, respectively, $\mathbf{q}$ is a vector lying in the first Brillouin zone (BZ), and δ is an infinitesimal number. Here and in the following, we consider the spin unpolarized case, that is, systems for which there is an equal number of electrons with spin up and down, each pair having the same spatial wave function.

The expression in Eq. (1) is rarely computed directly in *GW* codes, since one may need to employ a very dense grid to accurately represent the polarizability in real space. A common choice, which naturally takes care of periodic boundary conditions in solids, is to use a plane-wave (PW) basis set for both the polarizability and orbitals.

In a PW basis, each cell-periodic Bloch state $u_{n\mathbf{k}}(\mathbf{r}) = e^{-i\mathbf{k}\cdot\mathbf{r}}\psi_{n\mathbf{k}}(\mathbf{r})$ is represented by a vector containing the linear expansion coefficients $c_{n\mathbf{k}}(\mathbf{G})$, where $\mathbf{G}$ is a reciprocal-lattice vector. In this basis, the polarizability defined in Eq. (1) can be written as

$$\chi^0_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega) = 2\sum_{\mathbf{k}}\sum_{n}^{N_v}\sum_{m}^{N_c}\left[M^{\mathbf{G}}_{nm\mathbf{k}}(\mathbf{q})\right]^*\Delta_{nm\mathbf{k}}(\mathbf{q},\omega)\,M^{\mathbf{G}'}_{nm\mathbf{k}}(\mathbf{q}), \quad (3)$$

where the frequency independent matrix elements $M^{\mathbf{G}}_{nm\mathbf{k}}(\mathbf{q})$ are defined as

$$M^{\mathbf{G}}_{nm\mathbf{k}}(\mathbf{q}) = \langle \psi_{n\mathbf{k}+\mathbf{q}} | e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} | \psi_{m\mathbf{k}} \rangle. \tag{4}$$

The size of the matrix $\chi^0_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega)$ is thus determined by the number of $\mathbf{G}$ vectors employed in the expansion, which is typically defined in terms of an energy cutoff E^{χ}_{cut} for the dielectric matrix by including a number N_G of $\mathbf{G}$ vectors such that $|\mathbf{G}+\mathbf{q}|^2 \leqslant E^{\chi}_{\text{cut}}$ (in Rydberg atomic units). In practice, N_G is about an order of magnitude smaller than the number N_ψ of $\mathbf{G}$ vectors necessary to expand the Bloch functions in reciprocal space. In order to obtain converged calculations, the number of empty states N_c included in Eq. (3) is typically of the same order as N_G.

Once the matrices $\chi^0_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega)$ are obtained, one can compute the dielectric matrix as

$$\varepsilon_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega) = \delta_{\mathbf{G}\mathbf{G}'} - v(\mathbf{q}+\mathbf{G})\chi^0_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega), \tag{5}$$

where $v(\mathbf{q}+\mathbf{G})$ is the bare Coulomb potential expressed in reciprocal space. One then readily obtains the screened Coulomb potential $W_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega)$ after inverting the dielectric matrix in Eq. (5) for all $\mathbf{q}$ points and frequencies ω,

$$W_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega) = \varepsilon^{-1}_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega) v(\mathbf{q}+\mathbf{G}'). \tag{6}$$

It is convenient to split the screened Coulomb interaction matrix into two components: one that includes only the bare Coulomb interaction, and another that only includes electron correlations due to electronic screening,

$$W^c_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega) = W_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega) - v(\mathbf{q}+\mathbf{G})\delta_{\mathbf{G}\mathbf{G}'}. \tag{7}$$

With the separation in Eq. (7), the self-energy matrix elements are split into correlation $\langle \psi_{i\mathbf{k}} | \Sigma^c(E) | \psi_{j\mathbf{k}} \rangle$ and bare-exchange $\langle \psi_{i\mathbf{k}} | \Sigma^x | \psi_{j\mathbf{k}} \rangle$ contributions. The latter, called exchange self-energy, is equivalent to the nonlocal Hartree-Fock exchange in hybrid DFT calculations [3,35–38]. The typical challenge in a *GW* code is to efficiently compute the correlation contribution to the electronic self-energy Σ^c (the relevant expressions of Σ^x Σ^c can be found in Refs. [3,39]).

Since the self-energy is a direct product in real space and time domains of $G(\mathbf{r},\mathbf{r}',t)$ and $W(\mathbf{r},\mathbf{r}',t)$, it involves an integration when written in frequency space. Fortunately, this integration can be performed in a numerically efficient way by using the contour deformation (CD) technique, wherein real-frequency integrals involving $W^c(\omega)$ and $G(\omega)$ are written in terms of an integral over the imaginary axis, where both $W^c(\omega)$ and $G(\omega)$ are smooth quantities, plus a small number of residuals on the real axis [4,30,39–41]. While this formalism still requires one to compute $\chi^0(\omega)$ on a number of frequencies typically ranging from a few tens up to a hundred, it requires many less evaluations of $\chi^0(\omega)$ than a direct integration on the real axis.

If one is interested in computing the *GW* quasiparticle energy for only a selected number of states, such as when computing the quasiparticle band gap, then the most time-consuming part of the calculation within a sum-over-bands approach is in the construction of the inverse dielectric matrix. The main computational steps for evaluating $\varepsilon^{-1}_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega)$ (for instance, in the BERKELEYGW [3] software package) are the following.

TABLE I. Computational cost and required memory for the evaluation of the inverse $\varepsilon(\mathbf{q},\omega)$ at each $\mathbf{q}$ point. In the table, N_ψ is the size of the PW basis used to expand the KS-DFT wave functions, N_G is the size of the PW basis used to expand the polarizability and dielectric matrices, N_c and N_v are the number of conduction and valence bands respectively, N_k is the number of symmetry-reduced $\mathbf{k}$ points using the symmetry subgroup that leaves the $\mathbf{q}$ vector invariant, and N_ω is the number of frequencies (both imaginary and real) employed in the calculation. Only N_v, N_c, N_ψ, and N_G scale with the system size, the number of $\mathbf{k}/\mathbf{q}$ points in general scale inversely with the system size, and N_ω only depends on the system type.

	Execution	Memory
Matrix element	$O(N_k N_v N_c N_\psi \log N_\psi)$	$O(N_k N_v N_c N_G)$
Polarizability	$O(N_\omega N_k N_v N_c N_G^2)$	$O(N_\omega N_G^2)$
Inversion	$O(N_\omega N_G^3)$	$O(N_\omega N_G^2)$
Input and output	$O(N_\omega N_G^2)$	$O(N_\omega N_G^2)$

(1) Calculation of the PW matrix elements $M^{\mathbf{G}}_{nm\mathbf{k}}(\mathbf{q})$, given by Eq. (4). Each matrix element can be written as a convolution and thus efficiently computed using fast Fourier transforms (FFTs). Each individual FFT requires a number of operations that scales as $O(N \log N)$, with N proportional to system size. For a single value of $\mathbf{q}$, the required number of FFTs is proportional to $N_v N_c N_k$, with both N_v and N_c growing linearly with system size, resulting in an overall scaling of $O(N^3 \log N)$. The set of all matrix elements for a given $\mathbf{q}$ are stored in a matrix $\mathbf{M}(\mathbf{q})$ with dimensions $N_v N_c N_k \times N_G$, with the set of $(n, m, \mathbf{k})$ indices labeling different rows of $\mathbf{M}(\mathbf{q})$, and with each $\mathbf{G}$ vector corresponding to a different column.

(2) Calculation of the frequency-dependent RPA polarizability according to Eq. (3), which can be cast into a compact matrix notation as

$$\boldsymbol{\chi}^0(\mathbf{q},\omega) = 2\mathbf{M}^{\dagger}(\mathbf{q})\,\boldsymbol{\Delta}(\mathbf{q},\omega)\,\mathbf{M}(\mathbf{q}). \tag{8}$$

Here, $\mathbf{M}(\mathbf{q})$ is the previously calculated rectangular matrix, and $\boldsymbol{\Delta}(\mathbf{q},\omega)$ is a diagonal matrix with elements $\Delta_{nm\mathbf{k}}(\mathbf{q},\omega)$, defined in Eq. (2). From a computational stand point, Eq. (8) is a matrix multiplication involving two "tall and skinny" matrices, which is implemented in a parallel algorithm by employing a tailored data layout [31]. This step represents the most computationally demanding part of the algorithm, scaling asymptotically as $O(N^4)$.

(3) Computation of the frequency-dependent dielectric matrix ε, defined in Eq. (5), and its inverse. These steps are accomplished as two simple algebraic operations, the most demanding of which is a matrix inversion. The size of the matrix ε is N_G, which is proportional to N, resulting in a computational cost scaling as $O(N^3)$.

These steps are performed for all $\mathbf{q}$ points given in the input, offering another level of parallelization achieved by splitting the set of $\mathbf{q}$ points and performing the computation independently for each subset. A summary of the computational complexity and memory usage with respect to the calculation parameters is reported in Table I.

In summary, the implementation presented so far computes the polarizability matrix directly in reciprocal and frequency spaces. Overall the algorithm display an $O(N_\omega N_k N_c N_v N_G^2)$ computational bottleneck (scaling as N^4 with system size),

typical of traditional *GW* calculations; thus, any representation of the polarizability matrix that allows one to decrease the size of the basis size N_G would therefore dramatically speed up FF *GW* calculations.

III. STATIC SUBSPACE APPROXIMATION OF THE FREQUENCY-DEPENDENT INVERSE DIELECTRIC MATRIX

In this section, we describe a flexible algorithm which allows one to employ an alternative basis to decrease the computational cost to construct the full frequency-dependent dielectric matrix. The advantage of the algorithm presented here is that it is of general applicability, as our implementation can use any type of alternative basis. In fact, there is a variety of possible low-rank approximations that can be used in *GW* calculations [30]. As we discuss later, this paves the way to combine low-rank approaches with cubic-scaling methods to build the polarizability matrix for full-frequency *GW* calculations.

The approach considered here, inspired by the earlier work of Wilson *et al.* [26,27], Nguyen and coworkers [28], and Pham *et al.* [29], and motivated by the properties of the dielectric function [21–23], is to express the frequency dependence of the inverse dielectric matrix in a low-rank approximation fashion by using selected N_b eigenvectors of the static dielectric matrix, $\varepsilon_{\mathbf{GG'}}(\omega = 0)$, having eigenvalues larger than a given threshold. While each individual eigenvector of $\varepsilon^{-1}(\mathbf{q}, \omega)$ might change considerably as a function of ω, it was noted that the basis set spanned by the eigenvectors associated with the larger eigenvalues is not very dependent on ω for typical values of interest [15,16,19,21,29].

In this way, at the expense of a single matrix diagonalization, the cost for evaluating $\chi^0(\omega)$ and $\varepsilon^{-1}(\omega)$ for $\omega \neq 0$ is accelerated by a factor proportional to $(N_G/N_b)^2$ and $(N_G/N_b)^3$, respectively. Since the computational cost for a full-frequency calculation is roughly equal to that of repeating N_ω times a static calculation, with typically N_ω of the order of 10 to 100, even a moderate reduction in $N_G/N_b \simeq 3$–5 allows us to perform a full-frequency calculation in roughly the same order of time as it takes to perform a static or PPM calculation. Additionally, one can control the overall accuracy of the method by a single parameter, namely, the threshold t_{eig} for the truncation of the eigenspectrum of $\varepsilon(\omega = 0)$. In fact, as will be shown later, we find a direct relation between the absolute error in the QP energies and t_{eig}. Moreover, the error resulting from the static subspace approximation is largely independent from the other calculation parameters such as the number of bands and the dielectric matrix cutoff.

According to the definition of the dielectric matrix given in Eq. (5), $\varepsilon^{-1}(\omega)$, $\varepsilon(\omega)$, and $v \cdot \chi^0(\omega)$ all have the same eigenvectors. For practical reason, it is more convenient to define a symmetrized susceptibility,

$$\overline{\chi}^0_{\mathbf{GG'}}(\mathbf{q}, \omega) \equiv v^{\frac{1}{2}}(\mathbf{q} + \mathbf{G})\, \chi^0_{\mathbf{GG'}}(\mathbf{q}, \omega)\, v^{\frac{1}{2}}(\mathbf{q} + \mathbf{G'}), \quad (9)$$

which is Hermitian for $\omega = 0$. All eigenvalues of the static dielectric matrix are real and greater than one [21,22] so that the eigenvalues of $-\overline{\chi}^0$ are all positive. The eigenvalue spectrum of ε and $-\overline{\chi}^0$ at $\mathbf{q} = \Gamma$ and $\omega = 0$ for silicon carbide (β-SiC) are reported in Fig. 1, which shows the fast decay of the eigenvalues to one and zero for ε and $-\overline{\chi}^0$, respectively [26].

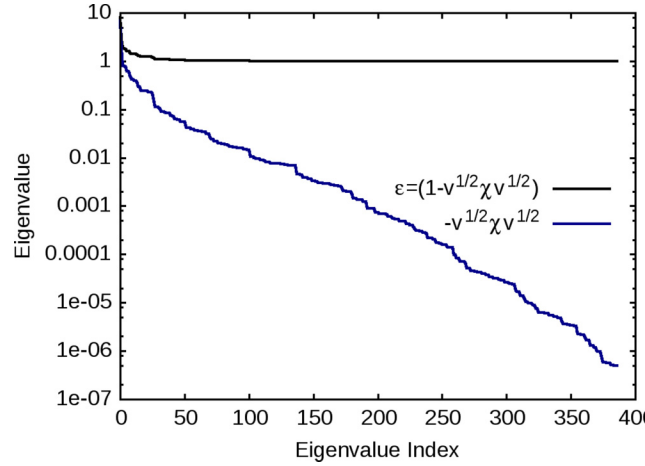

FIG. 1. Eigenvalue spectrum (logarithmic scale) for the static dielectric and symmetrized susceptibility matrices for β-SiC at Γ.

With the notation introduced so far and employing the same input parameters introduced in the previous section, the procedure for the computation of the frequency-dependent inverse symmetrized dielectric matrix within the static subspace approximation can be summarized as follows.

(1) For each $\mathbf{q}$ point, calculate $\chi^0_{\mathbf{GG'}}(\mathbf{q}, \omega = 0)$ using the standard procedure described in Section II. This step thus involves the computation of the matrix elements $M^{\mathbf{G}}_{nm\mathbf{k}}(\mathbf{q})$ and their contraction according to Eq. (8).

(2) Construct the symmetrized susceptibility $\overline{\chi}^0_{\mathbf{GG'}}$ according to the definition in Eq. (9), and perform the eigendecomposition

$$\overline{\boldsymbol{\chi}}^0(\mathbf{q}, 0) = \mathbf{C}(\mathbf{q})\, \boldsymbol{\lambda}(\mathbf{q})\, \mathbf{C}^\dagger(\mathbf{q}), \quad (10)$$

where $\mathbf{C}(\mathbf{q})$ is the $N_G \times N_G$ eigenvector matrix of $\overline{\chi}^0_{\mathbf{GG'}}(\mathbf{q}, 0)$, and $\boldsymbol{\lambda}(\mathbf{q})$ is the corresponding diagonal matrix of eigenvalues. Given a truncation threshold t_{eig}, define $\mathbf{C}_s(\mathbf{q})$ as the $N_G \times N_b$ truncated eigenvector matrix associated with the N_b absolute larger eigenvalues. This matrix is used to construct the low-rank approximation of $\overline{\boldsymbol{\chi}}^0$ and related quantities.

(3) For all other frequencies $\{\omega_i \neq 0\}$, the symmetrized susceptibility is projected onto the subspace defined by $\mathbf{C}_s$,

$$\overline{\boldsymbol{\chi}}^0_s(\mathbf{q}, \omega_i) = \mathbf{C}^\dagger_s(\mathbf{q})\, \overline{\boldsymbol{\chi}}^0(\mathbf{q}, \omega_i)\, \mathbf{C}_s(\mathbf{q}),$$

where $\overline{\boldsymbol{\chi}}^0(\mathbf{q}, \omega_i)$ and $\overline{\boldsymbol{\chi}}^0_s(\mathbf{q}, \omega_i)$ are the $N_G \times N_G$ and $N_b \times N_b$ matrix representations of the symmetrized susceptibility in the PW and static eigenvector basis, respectively. In order to take advantage of the reduced size of the static eigenvector basis, $\overline{\boldsymbol{\chi}}^0_s(\mathbf{q}, \omega_i)$ is computed in two steps. First, the PW matrix elements $M^{\mathbf{G}}_{nm\mathbf{k}}(\mathbf{q})$, which are frequency independent, are projected onto the $\overline{\boldsymbol{\chi}}^0_s$ subspace and scaled by the square root of the Coulomb potential,

$$\overline{\mathbf{M}}_s(\mathbf{q}) \equiv \mathbf{M}(\mathbf{q})\, \mathbf{v}^{1/2}(\mathbf{q})\, \mathbf{C}_s(\mathbf{q}), \quad (11)$$

TABLE II. Computational cost and required memory for the evaluation of the frequency-dependent inverse dielectric matrix within the static subspace approximation. The meaning of the symbols are the same as that given in Table I, with the addition of N_b defining the reduced number of eigenvectors of the static symmetrized susceptibility employed for the representation of the frequency-dependent part.

	Execution	Memory
Matrix elements	$O(N_k N_v N_c N_\psi \log N_\psi)$	$O(N_k N_v N_c N_G)$
Polarizability $\omega = 0$	$O(N_k N_v N_c N_G^2)$	$O(N_G^2)$
Eigendecomposition: $\mathbf{C}_s$	$O(N_G^3)$	$O(N_G^2)$
Basis transformation: $\overline{\mathbf{M}}_s$	$O(N_v N_c N_G N_b)$	$O(N_v N_c N_b)$
Polarizability $\omega \neq 0$	$O(N_\omega N_k N_v N_c N_b^2)$	$O(N_\omega N_b^2)$
Inversion	$O(N_\omega N_b^3)$	$O(N_\omega N_b^2)$
Input and output	$O(N_G N_b + N_\omega N_b^2)$	$O(N_G N_b + N_\omega N_b^2)$

where $\mathbf{v}^{1/2}(\mathbf{q})$ is an $N_G \times N_G$ diagonal matrix containing the Coulomb potential, $\mathbf{M}(\mathbf{q})$ is an $N_v N_c N_k \times N_G$ matrix, and $\overline{\mathbf{M}}_s(\mathbf{q})$ is an $N_v N_c N_k \times N_b$ matrix containing the subspace-projected PW matrix elements scaled by the square root of the Coulomb potential. We then compute $\overline{\boldsymbol{\chi}}_s^0(\mathbf{q}, \omega_i)$ as

$$\overline{\boldsymbol{\chi}}_s^0(\mathbf{q}, \omega_i) = \overline{\mathbf{M}}_s^\dagger(\mathbf{q})\, \boldsymbol{\Delta}(\mathbf{q}, \omega)\, \overline{\mathbf{M}}_s(\mathbf{q}), \tag{12}$$

which is very similar to Eq. (8) used in the previous algorithm. However, the computational cost associated with Eqs. (11) and (12) is $O(N_k N_v N_c N_G N_b)$ and $O(N_\omega N_k N_v N_c N_b^2)$, respectively, so that, compared to the $O(N_\omega N_k N_v N_c N_G^2)$ cost of the standard procedure, the computational effort gets reduced by a factor proportional to $(N_G/N_b)^2$, for large numbers of frequencies.

(4) The final step of the algorithm consists of the evaluation of the inverse dielectric matrix, which is also performed in the truncated basis. In particular, the $N_b \times N_b$ symmetrized dielectric matrix in the static eigenvector basis is simply obtained as $\bar{\boldsymbol{\varepsilon}}_s(\mathbf{q}, \omega_i) = \mathbf{I} - \overline{\boldsymbol{\chi}}_s^0(\mathbf{q}, \omega_i)$, which is then inverted numerically. Then, $\bar{\boldsymbol{\varepsilon}}_s^{-1}(\mathbf{q}, \omega_i)$ can either be used directly in this low-rank basis, or transformed back to the PW basis via

$$\bar{\boldsymbol{\varepsilon}}^{-1}(\mathbf{q}, \omega_i) = \mathbf{C}_s(\mathbf{q})\big[\bar{\boldsymbol{\varepsilon}}_s^{-1}(\mathbf{q}, \omega_i) - \mathbf{I}_{N_b}\big]\mathbf{C}_s^\dagger(\mathbf{q}) + \mathbf{I}_{N_G},$$

where $\mathbf{I}_{N_b}$ and $\mathbf{I}_{N_G}$ are identity matrices of size N_b and N_G, respectively. The reduced computational cost compared to the standard algorithm is cubic in the inversion step, namely, $O(N_b^3)$ instead of $O(N_G^3)$ and quadratic in the number of input and output operations. The unsymmetrized dielectric matrix in Eq. (6) is obtained as $\boldsymbol{\varepsilon}^{-1}(\mathbf{q}, \omega_i) = \mathbf{v}^{1/2}(\mathbf{q})\, \bar{\boldsymbol{\varepsilon}}^{-1}(\mathbf{q}, \omega_i)\, \mathbf{v}^{-1/2}(\mathbf{q})$.

The computational costs associated with the individual steps of the new procedure are reported in Table II. In summary, the algorithm described in this section addresses the memory and computational bottlenecks in full-frequency *GW* calculations by representing the frequency-dependent inverse dielectric matrix with the basis formed by the lower N_b eigenvectors of the symmetrized susceptibility, and which decreases the computational cost to compute the full-frequency dielectric matrix by $O(N_G/N_b)^2$ with N_G being the size of the original PW basis. Additionally, the inversion and storage of large $N_G \times N_G$ matrices are also avoided by reformulating the required operations in the reduced static N_b basis. The size of the reduced basis is determined by a single truncation parameter on the eigenspectrum of the static symmetrized susceptibility, namely t_{eig}, that is directly related to the overall accuracy of the resulting quasiparticle energies, as described in the next section.

The idea of speeding up *GW* calculation by employing a basis set obtained from the eigenvectors of the static dielectric matrix has been explored before, in particular in the context of using DFPT to avoid writing the Green's function in its spectral representation which requires a summation over empty states [4,25–29]. In contrast, the implementation presented here concerns *GW* formalisms that do not use density-functional perturbation theory and, as we will show in the next section, also enables one to significantly speed up FF *GW* calculations

Even though spectral representation and DFPT based *GW* calculations are often implemented in such a way as to have the same computational complexity with system size, the technical difficulty in obtaining many unoccupied KS orbitals for *GW* calculations is often cited as an advantage of DFPT-based approaches. However, there are still a number of possible algorithmic improvements that can be taken advantage of in formalisms that rely on the spectral representation of the Green's function. For instance, the explicit use of empty states allows the computation of the static polarizability with cubic computational effort [17], which can directly be incorporated into the present algorithmic framework.

In addition, for very large systems, extracting a large number of eigenvectors of the polarizability matrix can be a challenging tasks using iterative methods—especially since the eigenvalues of $\overline{\boldsymbol{\chi}}^0$ become clustered. We found this to be the case for system defects in semiconductors, which require large supercells containing over 1000 atoms, and when more than ∼10 000 eigenvectors of $\overline{\boldsymbol{\chi}}^0$ are necessary to converge absolute quasiparticle energies to within ∼50 meV. In these cases, having an algorithm based on direct methods—which is used in the implementation we propose here—may be considerably more stable.

IV. BENCHMARK CALCULATIONS

In this section, we present a series of benchmark calculations performed to validate the method and to assess how the error introduced by the static subspace approximation depends on the various input parameters as well as a function of the system type under study. The analysis here is organized in subsections each reporting the results for a particular class of materials, for which different approximations and computational strategies have to be considered. The studied materials include semiconductors, metals, systems with reduced dimensionality and molecules. Unless otherwise stated, QUANTUM ESPRESSO [42] and BERKELEYGW [3] have been used to perform the DFT and *GW* calculations, respectively. At the DFT level the calculations have been performed using norm-conserving pseudopotentials [43] and plane-wave basis. A detailed description of the computational protocol employed in these calculations is reported in the Supporting Information (SI) [44].

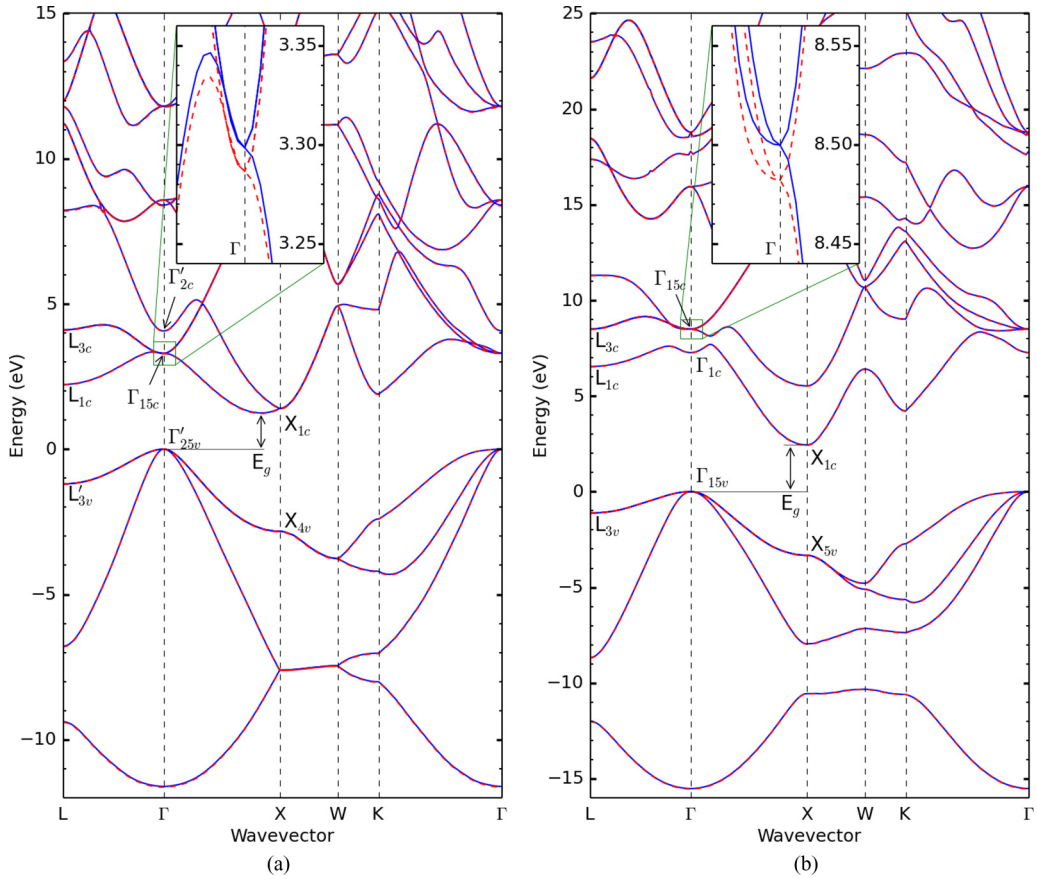

FIG. 2. Band structures of (a) silicon and (b) β-SiC, along with the error introduced by the static subspace approximation. The solid blue line is obtained without approximation and the red dashed employs the static subspace approximation with an eigenvalue screening threshold $t_{\text{eig}} = 0.01$. The zero reference in both cases has been set to the valence band maximum (VBM). The inset blow-ups show the band structure around the Γ_{15c} point to show the difference between the reference and approximate results.

A. Semiconductors

As a prototype for condensed phase semiconductor systems, we chose silicon (Si) and silicon carbide (SiC). These systems are among the most heavily studied semiconductors and this choice allows us to carefully assess our results compared to several previously reported calculations. We considered the cubic phase of silicon carbide (3C-SiC) also referred to as β-SiC. The experimental lattice parameter has been used for both silicon (5.43 Å) and SiC (4.36 Å). More details about the computational setups can be found in the SI [44].

The results obtained with the static subspace approximation compared to the reference calculation (without approximation) are summarized in Figs. 2 and 3. In particular, Figs. 2(a) and 2(b) show the quasiparticle band structures as obtained with and without the approximation, in both cases the zero is set to the valence band maximum (VBM). Even using a relatively high screening threshold for the eigenvalues (in this case $t_{\text{eig}} = 0.01$), the plots show excellent agreement over the whole range of considered energies and wave vectors. The inset in both cases shows a blow-up of the band structure around a specific region in order to highlight the deviation between reference and approximate results. By analyzing the error as signed deviation, i.e., taking the actual difference between each approximate/reference pair of quasiparticle energies calculated, we observed that for both systems considered, for a given t_{eig} the error is of similar magnitude over the whole range of energies and the approximate results approach the reference value from above by reducing t_{eig} (see SI [44] for more details). The mean absolute error (MAE) over all calculated quasiparticle energies is reported for silicon in Fig. 3 as a function of t_{eig}. The corresponding analysis for SiC is reported in SI [44]. In the plots, the label "relative" refers to the case for which the quasiparticle energies have been shifted with respect to the Fermi level, whereas "absolute" refers to the error calculated for the unshifted quasiparticle energies. As shown in the plots the error converges much faster with respect to t_{eig} for the relative energy than for the absolute energy. This is related to the fact that the static subspace approximation (see also SI [44]) introduces a uniform error over the different quasiparticles corrections that is compensated when calculating energy differences. These results show that excellent approximate *GW* solutions can be obtained using a screening threshold on the eigenvalues between 10^{-3} to 10^{-2}, and that the error can be further reduced by decreasing the value of t_{eig}.

Figure 4 presents the computational effort of our calculation as a function of the eigenvalue screening threshold (t_{eig}). The computational savings are measured with the reduction in the time to solution (percentage of the reference calculation).

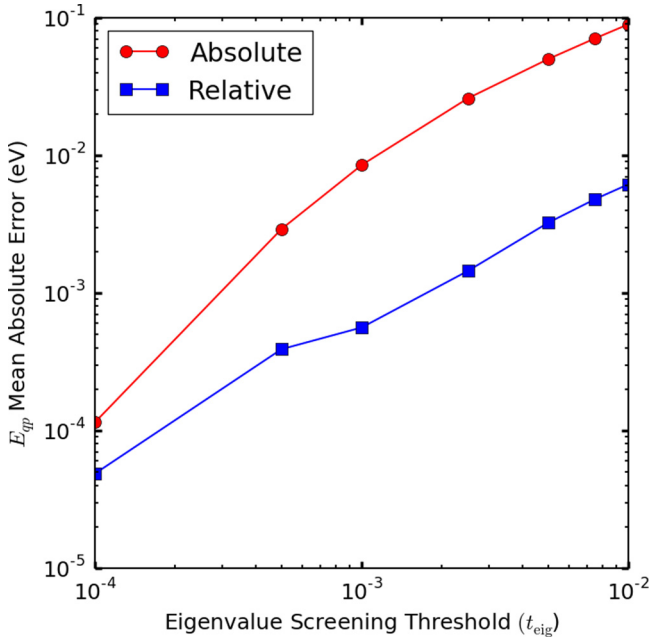

FIG. 3. For silicon the mean absolute error between the reference and approximate results calculated for 196 quasiparticle energies is reported as a function of t_{eig}, in this case, the error is reported with (relative) and without (absolute) shift with respect to the Fermi level.

The reduction in the number of eigenvectors, that is, the size of the subspace basis, is also given as a function of t_{eig}. As shown in the plots, the full-frequency calculation can be sped up by a factor ranging from four to twenty by choosing t_{eig} between 10^{-3} and 10^{-2}. This shows that, using the static subspace approximation, it is possible to control the error in a systematic way and perform full-frequency calculations with a computational effort being of the same order as using a PPM, for which only the static inverse dielectric matrix is computed.

Tables III and IV report the quasiparticle energies evaluated at high-symmetry points for silicon and β-SiC, respectively. Additionally Table V gives the analogous results calculated for aluminum arsenide (AlAs) obtained with the same procedure described in SI [44]. Note in this case that the wave functions employed in the evaluation of the self-energy have been generated over **k** grids centered at each of the high-symmetry points and both LDA and PBE [50] have been considered as a DFT starting point for G_0W_0. For comparison, the tables have been supplemented with available experimental and calculated data, the latter reported for a comparable level of theory. In general good agreement, within a few hundreds meVs, is achieved for silicon and AlAs while larger deviations are observed for SiC, especially when moving away from the valence band maximum.

Finally the full-frequency dependence of the self-energy, real $\mathrm{Re}\Sigma_{n\mathbf{k}}(\omega)$ and imaginary $\mathrm{Im}\Sigma_{n\mathbf{k}}(\omega)$ parts, are reported in Fig. 5 for the first eight bands of silicon. These quantities can be used to construct the spectral function [3] (see also SI [44]), which can be compared directly with photoemission spectra and other experimental band structure parameters. The possibility to calculate such quantities over a large range of frequencies requires, within the contour deformation formalism, the evaluation of the inverse dielectric matrix over a similarly large grid of frequencies on the real axis, which is made particularly inexpensive by using the static subspace approximation. The results reported in Fig. 5, obtained by employing an eigenvalues screening threshold of 10^{-2}, are in excellent agreement with previously reported calculation [39,51].

B. Metals: Copper

As the next benchmark, we focus our analysis on a metallic system, bulk copper. In contrast to insulators and semiconductors, metallic systems require additional considerations in the evaluation of the dielectric matrix due to the absence of a gap between occupied and empty states, leading to the possibility for intra-band transitions, that is, excitation within the same band. A fine **k**-point grid must therefore be used to accurately sample the possible transitions across the Fermi surface [57], as described in SI [44].

Due to the large number of **q**/**k** points involved in the calculations, which has a large impact on the eigenspectrum

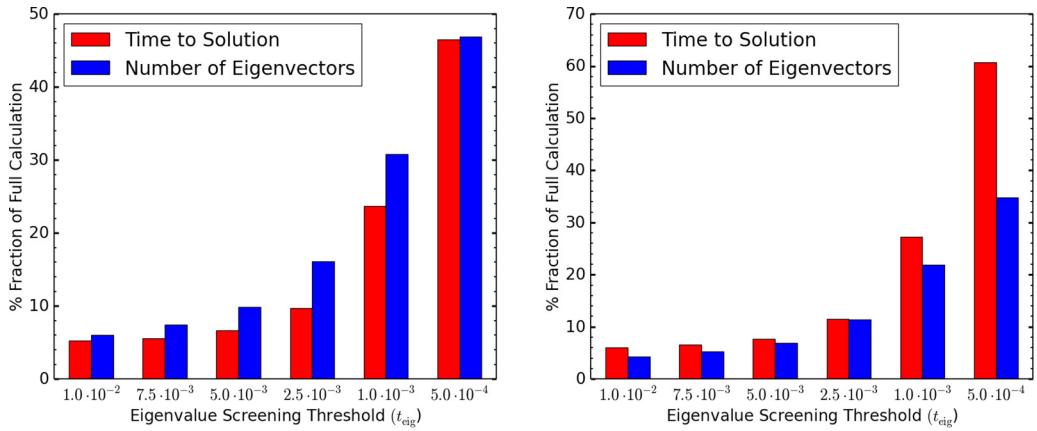

FIG. 4. For silicon (a) and β-SiC (b), blue columns show the number of eigenvectors in percentage of the total that are retained in the computation of the inverse dielectric matrix (ε) as a function of the eigenvalue screening threshold (t_{eig}). Red columns report the percentage reduction of the time to solution for the evaluation of ε with respect to the reference calculation (no approximation).

TABLE III. G_0W_0 quasiparticle energies calculated at high-symmetry points for silicon, in eV, relative to the VBM. We employed an eigenvalue threshold of $t_{\text{eig}} = 10^{-3}$. The second line reports the mean-field DFT functional employed as the starting point for the *GW* calculation. All theoretical calculations use a full-frequency treatment of the dielectric matrix. Experimental values reported as quoted in Ref. [45].

	This Work		Ref. [45]	Ref. [46]	Ref. [29]	Ref. [4]	Ref. [47]	
	LDA	PBE	LDA	LDA	LDA	PBE	PBE	Exp.
Γ_{1v}	−11.54	−11.61	−11.57	−11.57	−11.64		−11.83	-12.5 ± 0.6
Γ'_{25v}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Γ_{15c}	3.32	3.30	3.24	3.23	3.25	3.32	3.25	3.40, 3.05
Γ'_{2c}	3.82	4.07	3.94	3.96	3.92			4.23, 4.1
X_{1v}	−7.53	−7.57	−7.67	−7.57	−7.75			
X_{4v}	−2.77	−2.83	−2.80	−2.83	−2.88	−2.96	−2.86	$-2.9, -3.3 \pm 0.2$
X_{1c}	1.46	1.40	1.34	1.35	1.36	1.37	1.28	1.25
X_{4c}	10.45	10.59	10.54					
L'_{2v}	−9.33	−9.38	−9.39	−9.35	−9.38			-9.3 ± 0.4
L_{1v}	−6.69	−6.76	−6.86	−6.78	−6.93			-6.7 ± 0.2
L'_{3v}	−1.18	−1.20	−1.17	−1.20	−1.23	−1.21	−1.21	$-1.2 \pm 0.2, -1.5$
L_{1c}	2.17	2.23	2.14	2.18	2.21	2.29	2.14	$2.1, 2.4 \pm 0.15$
L_{3c}	4.17	4.11	4.05	4.06	4.00			4.15 ± 0.1

of the symmetrized susceptibility, the static subspace approximation has been tested by fixing the number of included eigenvectors rather than selecting them by using a screening threshold on the eigenvalues. This choice is more reasonable and results in better systematic convergence in metals compared to the case of semiconductors. The error introduced by the static subspace approximation has been tested by performing the calculation of ten quasiparticle energies (roughly five above and below the Fermi level depending on the **k** point considered) for a set of five **k** points, including Γ, X, L, and two additional points in the vicinity of the Fermi surface. Figure 6 reports the convergence of the mean absolute error as a function of the number of eigenvectors included in the subspace basis (ranging between 10%–30% of the total). As in the previous section, the error is reported as relative and absolute to distinguish between the case for which the quasiparticle energies have and have not been shifted with respect to the Fermi level respectively. The plot shows a systematic reduction of the error introduced by the static subspace approximation by increasing the subspace basis and that mean error below 20 meV can be achieved by using less than 30% of the total number of eigenvectors.

Figure 7 reports the PBE and *GW* quasiparticle band structures computed employing a cutoff of 45 Ry and employing 100 eigenvectors in the subspace basis. Details of the energy level positioning at high-symmetry points are reported in Table VI compared with experiments and previously reported calculations.

TABLE IV. G_0W_0 quasiparticle energies calculated at high-symmetry points for β-SiC, in eV, relative to the VBM. We employ an eigenvalue threshold $t_{\text{eig}} = 10^{-3}$. The second line reports the mean-field DFT functional employed as the starting point for the *GW* calculation. PPM indicates that the calculations have been obtained within a plasmon-pole model. The other theoretical calculations use a full-frequency treatment of the dielectric matrix. Experimental values reported as quoted in Ref. [48].

	This Work		Ref. [29]	Ref. [4]	Ref.[47]	Ref. [48]	Ref. [49]	
	LDA	PBE	LDA	PBE	PBE	LDA-PPM	LDA-PPM	Exp.
Γ_{1v}	−15.41	−15.52	−15.54		−15.69	−16.08	−16.44	
Γ'_{15v}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
Γ_{1c}	7.31	7.27	7.26	7.52	7.35	7.19	7.35	7.4
Γ_{15c}	8.45	8.50	8.10			8.18	8.35	7.75
X_{1v}	−10.37	−10.58	−10.46			−10.96	−11.24	
X_{3v}	−7.80	−7.87	−8.17			−8.44	−8.64	
X_{5v}	−3.28	−3.33	−3.47	−3.46	−3.30	−3.53	−3.62	−3.4
X_{1c}	2.57	2.42	2.31	2.28	2.42	2.19	2.34	2.39, 2.42
X_{3c}	5.54	5.52	5.41			5.23	5.59	5.2(3)
L_{1v}	−11.85	−12.03	−12.06			−12.46	−12.75	
L_{1v}	−8.53	−8.64	−8.92			−9.19	−9.42	
L_{3v}	−1.12	−1.12	−1.10	−1.16	−1.10	−1.21	−1.21	−1.15
L_{1c}	6.53	6.54	6.43	6.37	6.62	6.30	6.53	6.35
L_{3c}	8.52	8.50	8.32			8.07	8.57	8.55

TABLE V. G_0W_0 quasiparticle energies calculated at high-symmetry points for AlAs, in eV, relative to the VBM. We employed an eigenvalue threshold $t_{\text{eig}} = 10^{-3}$. The second line reports the mean-field DFT functional employed as a starting point for the *GW* calculation. PPM indicates that the calculations have been obtained within a plasmon-pole model. The other theoretical calculations use a full-frequency treatment of the dielectric matrix. Experimental values reported as quoted in Ref. [48].

	This Work		Ref. [29]	Ref. [4]	Ref. [47]	Ref. [48]	
	LDA	PBE	LDA	PBE	PBE	LDA-PPM	Exp.
Γ_{1v}	−11.68	−11.70	−11.66		−11.82	−11.51	
Γ'_{15v}	0.00	0.00	0.00	0.00	0.00	0.00	
Γ_{1c}	2.90	2.99	2.96	2.99	2.99	2.74	3.13
Γ_{15c}	5.13	5.07	5.07			5.06	
X_{1v}	−9.75	−9.70	−9.77			−9.67	
X_{2v}	−5.31	−5.37	−5.37			−5.55	
X_{5v}	−2.18	−2.22	−2.20	−2.35	−2.17	−2.27	−2.41
X_{1c}	2.27	2.14	2.13	2.01	2.31	2.16	2.23
X_{3c}	3.12	3.03	3.08			3.04	
L_{1v}	−10.31	−10.28	−10.27			−10.19	
L_{1v}	−5.44	−5.52	−5.82			−5.69	
L_{3v}	−0.84	−0.86	−0.90	−0.90	−0.82	−0.87	
L_{1c}	2.96	2.96	3.02	2.94	3.08	2.84	2.36
L_{1c}	5.58	5.49	5.63			5.52	

Compared to experiments, the calculated positioning of the *d* bands are systematically (∼0.6 eV) above the experimental values while the widths, with the exception of $L_3 - L_1$, are very well reproduced. Opposite is the case of the *s*/*p* bands which show a reversed trend, that is systematically below experiments. These results are consistent with previous studies [58,59] showing that spurious self-interactions in the *GW* approximation may result in upward shifts of highly localized *d* states up to 0.6 eV and downward shifts for dispersed states at the valence or conduction band edges up to 0.3 eV.

Compared to other theoretical results, our calculations are in good agreement with those reported by Liu *et al.* [17], both for the level positioning and widths. Additionally, compared to the results reported in Ref. [17], we also found good agreement in the overall profile of the band structure (see Fig. 7) and for the computed spectral function of the Green's function at Γ (see SI [44]).

C. Systems with reduced dimensionality: Monolayer MoS_2

Another class of materials that is computationally demanding for the *ab initio GW* approach is that of systems with reduced dimensionality, which includes molecules, clusters, nanotubes, nanoribbons, slabs, and monolayer materials. The two main challenges associated with these materials is that (1) when performing calculations in a PW basis, we need to construct large enough supercells and truncate the Coulomb potential to avoid spurious interactions between the repeated cell [60], and (2) systems with reduced dimensionality often display a strong spatial variation of the dielectric screening, which manifests in a slow convergence of these calculations with **k**-point sampling.

To assess the validity of our approximation in these systems, we perform a thorough analysis on the quasiparticle properties of monolayer MoS_2, a prototypical quasi-two-dimensional (quasi-2D) semiconductor, which belongs to the

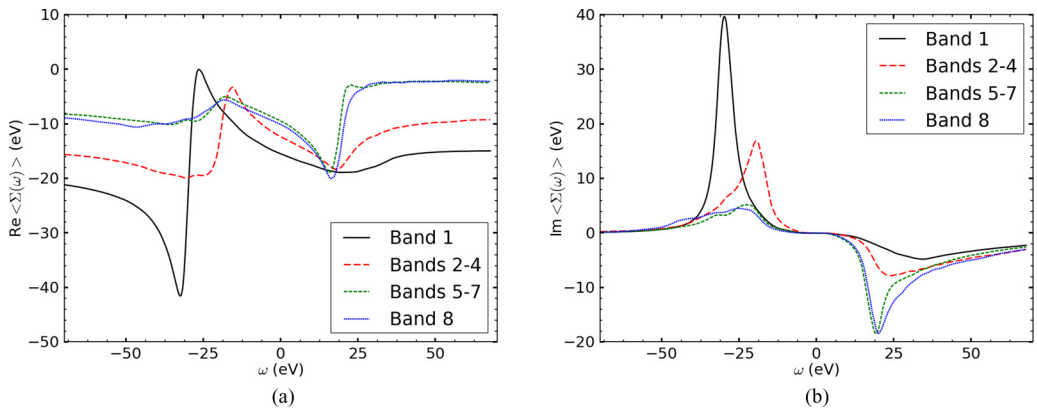

FIG. 5. The frequency dependence of the matrix elements of the time-ordered self-energy operator for the first eight bands of silicon evaluated at the Γ point. The zero of the frequency axis is set to the center of the gap. (a) and (b) show the real and imaginary part of $\langle\psi_{n\mathbf{k}}|\Sigma(\omega)|\psi_{n\mathbf{k}}\rangle$, respectively. The reported calculations are in excellent agreement with previous works [39,51].

TABLE VI. Quasiparticle energies calculated at high-symmetry points for copper, in eV, relative to the Fermi level. Experimental values taken from Ref. [56]. For calculation parameters see Ref. [44].

		PBE	G_0W_0	Ref. [17]	Ref. [55]	Exp.
Position d bands	Γ_{12}	−2.21	−2.16	−2.11	−2.81	−2.78
	X_5	−1.49	−1.42	−1.45	−2.04	−2.01
	L_3	−1.63	−1.57	−1.58	−2.24	−2.25
	$\Gamma_{12}-\Gamma_{25'}$	0.85	0.72	0.69	0.60	0.81
Widths of d bands	X_5-X_3	2.98	2.73	2.60	2.49	2.79
	X_5-X_1	3.43	3.20	3.10	2.90	3.17
	L_3-L_3	1.45	1.36	1.26	1.26	1.37
	L_3-L_1	3.52	3.29	3.16	2.83	2.91
Position of s/p bands	Γ_1	−9.44	−9.45	−9.18	−9.24	−8.60
	$L_{2'}$	−1.04	−1.14	−1.02	−0.57	−0.85
L gap	$L_{1c}-L_{2'}$	4.73	4.99	4.98	4.76	4.95

family of transition metal dichalcogenides. This system displays several features which make *GW* calculations computationally demanding. Because the VBM and CBM of this material have a considerable amount of 4*d* character, it is crucial to include the exchange interaction originating from semicore 4*s* and 4*p* states to accurately compute the quasiparticle band gap. This also reflects in a relatively large dielectric cutoff of ∼35 Ry being necessary to converge *GW* calculations for these systems. In addition, because of the reduced dimensionality, the dielectric function changes rapidly in the $\mathbf{q} \to 0$ limit, [61,62] which require very fine **k**-point sampling to converge *GW* calculations [63].

To address the slow convergence with respect to **k**-point sampling, we employ the recently developed nonuniform neck subsampling (NNS) method [64], which efficiently captures the sharp features of the dielectric matrix, by evaluating $\varepsilon^{-1}(\mathbf{q})$ on a nonuniform **q** grid. With the exception of the

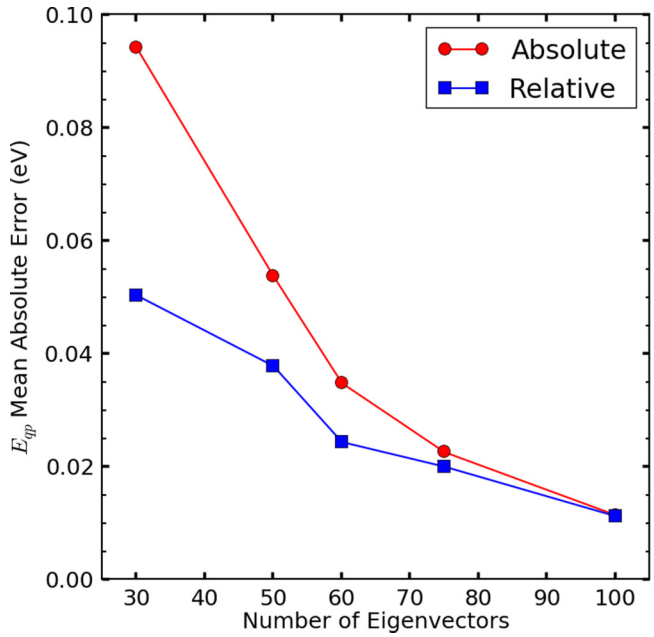

FIG. 6. Mean absolute error between the reference and approximate results for a total of 50 quasiparticle energies as a function of the number of eigenvectors included in the subspace basis. The error is reported with (relative) and without (absolute) shift with respect to the Fermi level. The calculations have been performed with a cutoff of 32 Ry and including 300 and 1000 bands in the calculation of the inverse dielectric matrix and self-energy, respectively. The considered number of eigenvectors ranges between 10%–30% of the full basis.

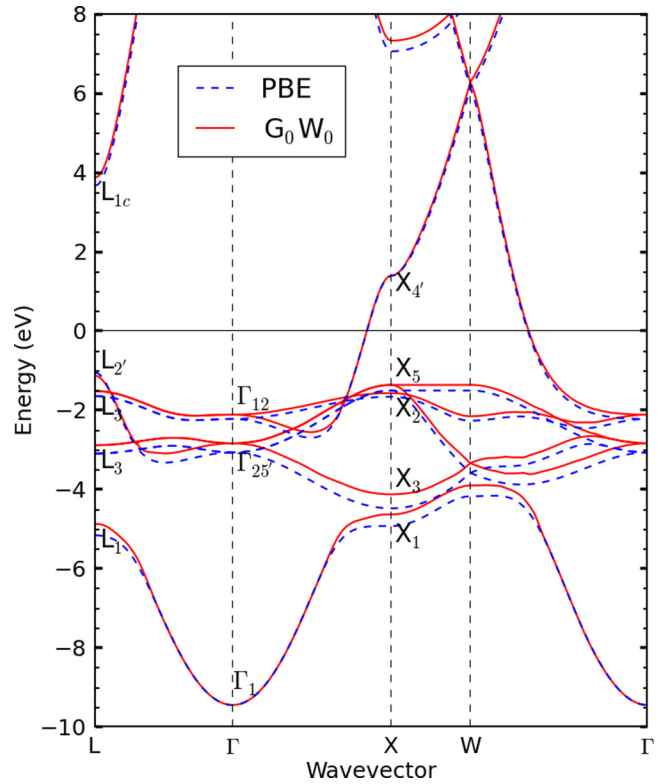

FIG. 7. Band structure for copper obtained by employing a cutoff of 45 Ry and including 400 and 1000 bands in the calculation of the inverse dielectric matrix and self-energy, respectively. The size of the subspace basis has been fixed to 100 eigenvectors (∼25% of the total). A uniform **k** grid [52] of $16 \times 16 \times 16$ and $32 \times 32 \times 32$ have been used for the evaluation of ε^{-1} for $\mathbf{q} \neq 0$ and $\mathbf{q} \to 0$ respectively. Calculations performed with a norm-conserving scalar-relativistic pseudopotential [53,54] including semicore *s* and *p* states.

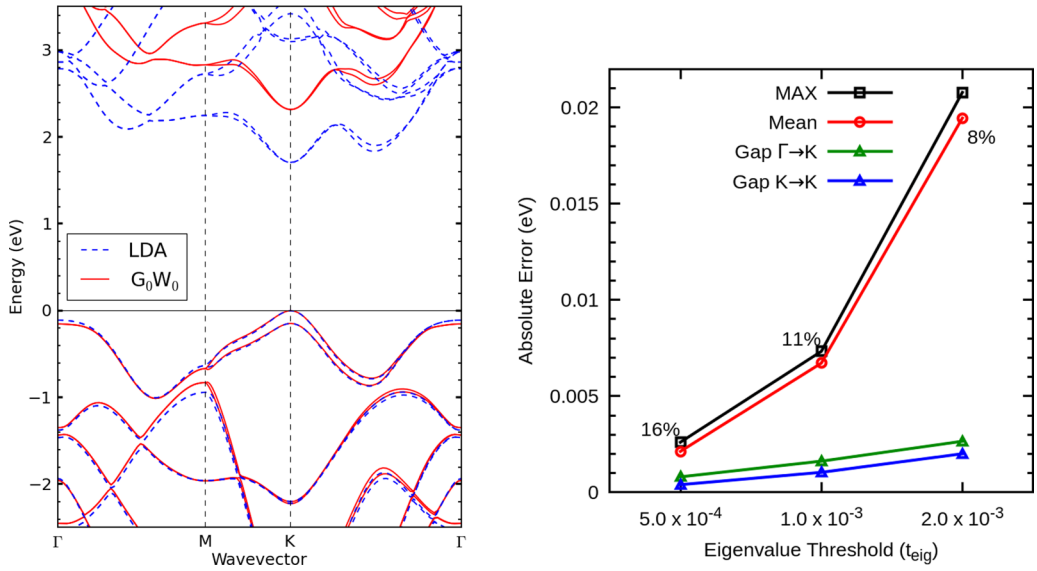

FIG. 8. Convergence study of monolayer MoS_2 with the static subspace approximation. (a) G_0W_0 quasiparticle band structure (red lines) of monolayer MoS_2 obtained with $N_b = 400$ eigenvectors, and the corresponding LDA band structure (blue dashed lines). (b) Convergence of the error of the quasiparticle evaluated at the highest valence and lowest conduction band states at K and Γ with respect to the screening threshold on the eigenvalues t_{eig}. We show the absolute error on the $\Gamma \to K$ and $K \to K$ band gaps (green and blue triangles respectively), the mean absolute error (red circles) and the maximum absolute error (black squares). The percentage labels in the plot shows the fraction of eigenvectors retained in the calculations for each t_{eig}.

utilization of the NNS technique, the remaining calculation parameters for the evaluation of the full-frequency quasiparticle corrections (cutoffs, number of bands *etc.*) are similar to those employed by Qiu *et al.* in Ref. [63], except that here we performed one-shot G_0W_0 calculations (i.e., without self-consistently updating the eigenvalues of the Green's function G). We note that these calculations have been performed on a large supercell and with a truncated Coulomb potential [60] to avoid spurious interactions between repeated supercells along the confined direction.

We show in Fig. 8(a) the calculated quasiparticle band structure of monolayer MoS_2, for which we used $N_b = 400$ eigenvectors ($\sim$12% of the total) in the generation of the static subspace. Spin-orbit interaction was included perturbatively following Ref. [61]. In addition, in Fig. 8(b), we report the error in the quasiparticle energies introduced by the static subspace approximation as a function of the eigenvalue screening threshold (t_{eig}) for both absolute and relative energy differences.

From Fig. 8(b), it is clear that we can obtain very good convergence of the quasiparticle energy of monolayer MoS_2 with an error of just 20 meV in the absolute value and a negligible error in the energy differences when we keep only 8% of the eigenvectors. It is interesting to note that, for similar dielectric cutoffs, the absolute number of eigenvectors N_b necessary to achieve a given target error for monolayer MoS_2 is similar to that needed in a bulk system such as copper. Hence, the fact that we only need to keep a much smaller fraction of the eigenvectors for monolayer MoS_2 is likely due to the presence of a large vacuum region, which increases the size of the epsilon matrix but brings little additional information on the dielectric function for the region inside the physical system of interest. We stress, however, that the presence of the vacuum may affect the dielectric environment outside the material, as discussed in Ref. [63]. These results show that the static subspace approximation is particularly useful for systems with reduced dimensionality. In fact, for the case of monolayer MoS_2, excellent convergence is achieved by using $\sim$10% of the original PW basis size, allowing one to speed up the evaluation of the inverse dielectric matrix by an order of magnitude.

D. Molecules: H_2O

As a final benchmark, we perform *GW* calculations on a water molecule to test the static subspace approach on a prototypical quasi-zero-dimensional system. Our *GW* calculations use a DFT mean-field starting point with the PBE functional, norm conserving pseudopotential [54,65] and cluster boundary conditions [66]. The details about the calculation are reported in the supporting information [44]; additionally, extensive analysis of convergence issues for isolated systems in PW-GW implementations (which goes beyond the scope of this work) can be found in Ref. [67]. Our calculated value of the HOMO quasiparticle energy is -11.84 eV, which compares well with previously reported calculations ranging from -11.8 to -12.1 eV [4,11,18,67,68].

For *GW* calculations that employ a traditional sum-over-bands approach, it is well known that the quasiparticle energies depend sensitively on the cutoff of the dielectric matrix and the number of bands included in the calculation, and that these two parameters are interconnected [69]. This is illustrated here for the case of a water molecule in Fig. 9. An important question therefore is if the static subspace approximation represents another convergence degree of freedom that must be studied simultaneously with

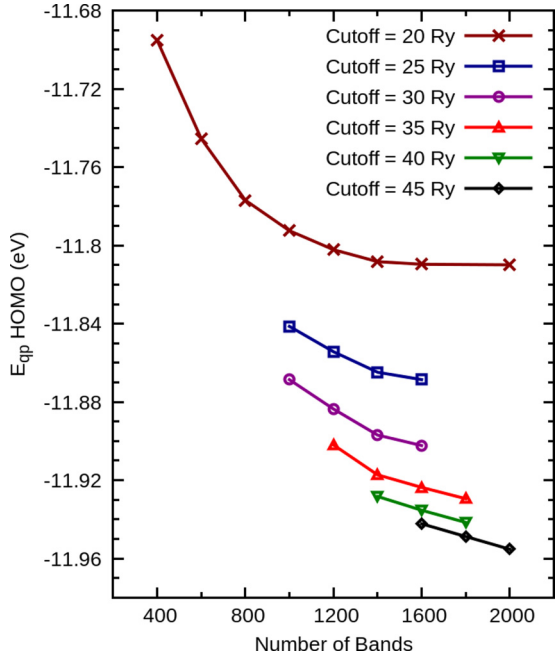

FIG. 9. Convergence of the HOMO quasiparticle energy of water with respect to the number of bands included in the calculation and the dielectric matrix ϵ and the screened Coulomb cutoff.

the cutoff of the dielectric matrix and the number of bands, or if they are independent. Fortunately, we show here that the subspace approximation is largely independent of other convergence parameters, which dramatically simplifies *GW* calculations within a sum-over-bands approach.

In Fig. 10(a), we show the convergence of the quasiparticle energies obtained for a fixed number of bands (1600), but varying the cutoff $E^{\varepsilon}_{\text{cut}}$ of the dielectric matrix, as well the eigenvalue threshold t_{eig} for the subspace calculations. We report the mean absolute error of the quasiparticle energy computed for four states for a given eigenvalue threshold and a given dielectric cutoff, relative to the calculation performed with the same cutoff but without the static subspace approximation. In Fig. 10(b), we report the complementary case, wherein we fix $E^{\varepsilon}_{\text{cut}} = 20$ Ry but vary the number of bands and t_{eig}, and report the error from the subspace approximation for a given number of bands. In all cases, even though the absolute quasiparticle energies change significantly with respect to dielectric cutoff and number of bands, the error introduced by the static subspace approximation as a function of t_{eig} displays very little dependence on the cutoff or number of bands employed. This implies that convergence studies with respect to t_{eig} can be dramatically simplified by performing calculations with relatively small cutoffs and number of bands to determine the appropriate value of t_{eig}. Once t_{eig} is determined for the desired accuracy of the calculation, one can then use standard techniques to converge the calculation with respect to the number of bands and dielectric cutoff, but taking advantage of the static subspace approximation to make these calculations more computationally efficient.

In all cases, less than 4% of the total number of eigenvectors is sufficient to provide excellent convergence in the absolute quasiparticle energy of $\sim 10^{-3}$ eV. This represents a reduction of the basis size that is even larger than that found for quasi-2D and bulk systems.

V. ZINC OXIDE

Despite the wide use and success of the *ab initio GW* approach to predict electronic properties of a large variety of condensed-phase systems, zinc oxide, an apparently simple insulator, turned out to be a challenging system for these first-principle calculations. Most of the difficulties arise from the presence of shallow Zn-3*d* states, which give strong covalent hybridization with valence O-2*p* states, and the presence of shallow semicore Zn-3*p* and 3*s* states, which also affect the valence band edges via exchange interactions. For the wurtzite structure, at the local and semilocal DFT levels, the electronic

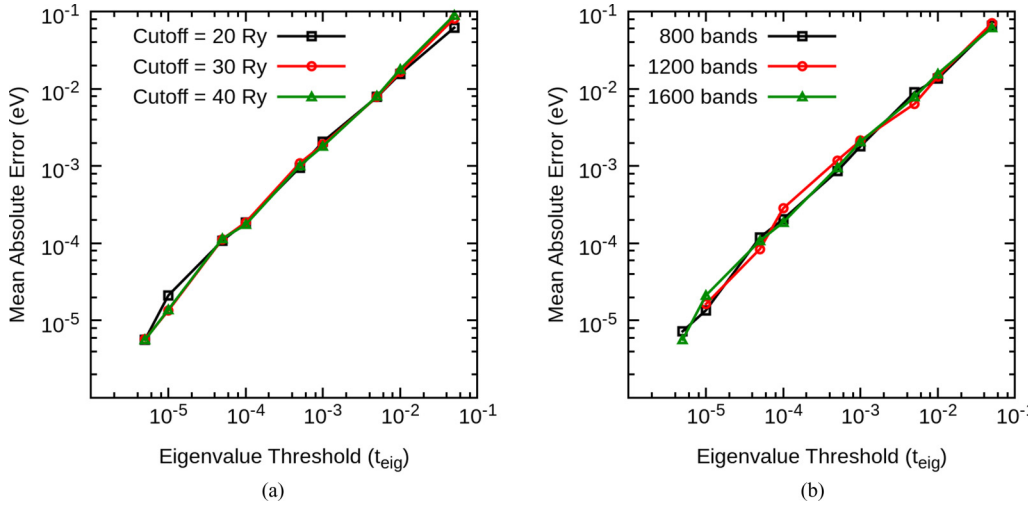

FIG. 10. (a) Convergence of the mean absolute error for the quasiparticle states with respect to the truncation threshold on the eigenvalues t_{eig}. Each curve was evaluated with a different cutoff for the dielectric matrix, but with a fixed number of bands (1600). The error is computed with respect of the calculation performed without the subspace approximation, but for the same cutoff of the dielectric matrix. (b) Similar plot as in (a), but where each curve was computed with a different number of bands, but with a fixed cutoff for the dielectric matrix of 20 Ry.

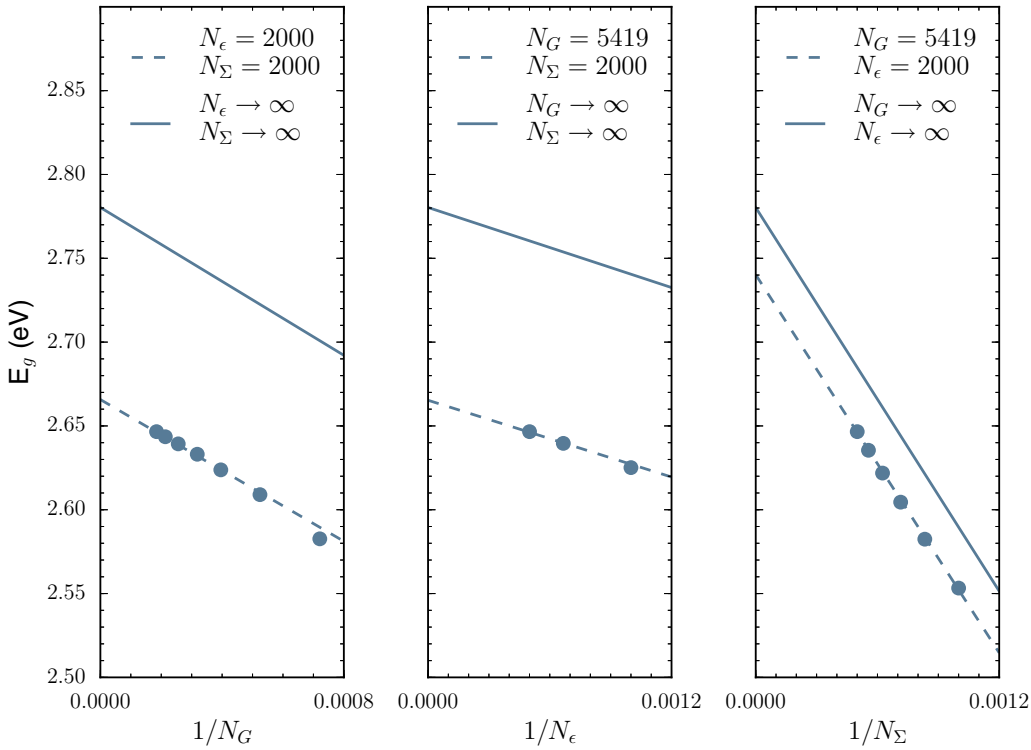

FIG. 11. Convergence study of the *GW* quasiparticle band gap of ZnO with respect to three interdependent parameters (see the text). The convergence function is plotted for two fixed parameters (dashed lines) used to produce the data points (discs), and for the extrapolated parameters (solid lines).

gap is strongly underestimated (0.7–0.9 eV) compared to experiments (3.6 eV [70,71]), while theoretical results at the *GW* level range between 2.3 and 4.5 eV [69,72–81]. Such a wide spread of results at the *GW* level, especially when considering the single-shot G_0W_0 method, is a direct consequence of the sensitivity of the approximation with respect to the various calculation parameters for this system. These include the poor starting point description for the electronic structure at the local and semilocal DFT levels, the importance of the core-valence exchange contributions, errors introduced by linearization procedures, plasmon-pole models, slow convergence with respect to the number of bands included in the evaluation of the polarizability and self-energy, and slow convergence with respect to the plane-wave cutoff for the expansion of the dielectric function. Therefore, because of its numerical sensitivity, ZnO is an ideal test system to benchmark the static subspace approximation.

In this section, we both validate the static subspace approximation in this computationally challenging system, and also take advantage of our algorithm to perform highly converged calculations on ZnO. Similarly to the case of the water molecule, we have verified that, for ZnO, the error resulting from the static subspace approximation is independent of the other convergence parameters. We also found that an eigenvalue screening threshold of $t_{\text{eig}} = 10^{-3}$ and a total of 81 frequencies on the real and imaginary axis is sufficient to converge absolute quasiparticle energies to within better than 10 meV. More details on these calculations are reported in the supporting information [44].

Apart from these quantities, there are four other convergence parameters that need to be considered in a *GW* calculation: the **k**-point sampling of the BZ, the size of the dielectric matrix in the PW basis (N_G), and the number of bands included in the evaluation of the dielectric function (N_ε) and self-energy (N_Σ). The first convergence parameter (k-point sampling) is in most cases independent on the other three, which are interdependent [82,83]. We perform a convergence study on all of these three parameters including up to 2000 bands in Σ and ε, with PW cutoffs up to 100 Ry (see SI [44] for the detailed analysis).

In order to extrapolate the quasiparticle band gap to the complete basis set limit ($N_G \to \infty$, $N_\varepsilon \to \infty$, $N_\Sigma \to \infty$), we make use of a convergence function that models the quasiparticle band gap as a function of the three parameters. Our convergence function is the product of three linear functions of $1/N_G$, $1/N_\varepsilon$, and $1/N_\Sigma$, and we fit the six coefficients to our data. Figure 11 shows a subset of our *GW* calculations used in our convergence study, as well as the convergence function for any two fixed parameters. The extrapolation to the complete basis set limit gives a ZnO quasiparticle band gap of $E_g = 2.78$ eV, which represents our best estimate of the fully converged gap within respect to all three parameters.

As a final note, we emphasize that the error introduced by the static subspace approximation in all examined cases (see SI [44]) is negligibly small—of the order of few meV in the absolute quasiparticle energies—while the approximation allows one to speed up these calculations by one order of magnitude. Therefore we expect the static subspace approximation

to be a valuable technique when dealing with computationally challenging systems such as ZnO.

VI. CONCLUSIONS

In this paper, we described a method—the static subspace approximation—that greatly speeds up the evaluation of the frequency-dependent part of the polarizability matrix for G_0W_0 calculations within the traditional sum-over-states approach. This is achieved by performing a low-rank approximation of the static ($\omega = 0$) symmetrized susceptibility matrix, and subsequently using the subspace formed by the eigenvectors to compress the relevant plane-wave matrix elements necessary to compute the polarizability matrix for other frequencies. This approximation is motivated by the properties of the eigenspectrum of the dielectric function [21–23] and is based on the earlier work of Wilson *et al.* [26,27], Nguyen and coworkers [28], and Pham *et al.* [29].

We tested the approximation for a wide variety of systems, showing that, depending on the dimensionality of the problem, retaining 5%–25% of the total number of eigenvectors is enough to obtain excellent accuracy in the evaluation of the final quasiparticle energies. In addition, we show that the error in the final quasiparticle energy can be directly controlled by setting a target eigenvalue threshold. We also obtain useful rules of thumb, wherein an absolute error of just a few meV's can be achieved by keeping eigenvalues up to a threshold of $\sim 10^{-3}$, while relative quasiparticle energies typically converge with a threshold of $\sim 10^{-2}$.

We also show that the error introduced by the static subspace approximation is largely independent of the error introduced by other convergence parameters, such as the number of bands used to represent the Green's function in its spectral form, and the plane-wave cutoff of the dielectric matrix. This is particularly important for converging *GW* calculations, since it shows that the error introduced by the subspace approximation can be determined with calculations using a relatively small cutoff for the dielectric matrix and including a smaller number of bands.

The approach we propose allows one to speed up the evaluation of G_0W_0 quasiparticle energies by at least one order of magnitude compared to the standard approaches, making the cost of a FF calculation of the same order as that based on a plasmon-pole model. This open the possibility not only to study large systems at the FF-*GW* level, but also to perform systematic convergence studies of complex materials. We show this by successfully applying the method to study the quasiparticle energies of zinc oxide, a system that is difficult to treat within the *ab initio GW* approach. The algorithm described in this paper has been implemented and is available in the BERKELEYGW software package [3].

ACKNOWLEDGMENTS

This work was supported by the Center for Computational Study of Excited-State Phenomena in Energy Materials (C2SEPEM) and by the SciDAC Program on Excited State Phenomena in Energy Materials at the Lawrence Berkeley National Laboratory, which is funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, Materials Sciences and Engineering Division under Contract No. DE-AC02-05CH11231, as part of the Computational Materials Sciences Program. This research used resources of the National Energy Research Scientific Computing Center (NERSC), 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.

[1] L. Hedin, Phys. Rev. **139**, A796 (1965).
[2] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986).
[3] J. Deslippe, G. Samsonidze, D. A. Strubbe, M. Jain, M. L. Cohen, and S. G. Louie, Comput. Phys. Commun. **183**, 1269 (2012).
[4] M. Govoni and G. Galli, J. Chem. Theory Comput. **11**, 2680 (2015).
[5] X. Gonze, G. Rignanese, M. Verstraete, J. Betiken, Y. Pouillon, R. Caracas, F. Jollet, M. Torrent, G. Zerah, M. Mikami, P. Ghosez, M. Veithen, J.-Y. Raty, V. Olevano, F. Bruneval, L. Reining, R. Godby, G. Onida, D. Hamann, and D. Allan, Z. Kristallogr. **220**, 558 (2005).
[6] H. Jiang, R. I. Gómez-Abal, X.-Z. Li, C. Meisenbichler, C. Ambrosch-Draxl, and M. Scheffler, Comp. Phys. Comm. **184**, 348 (2013).
[7] M. Shishkin and G. Kresse, Phys. Rev. B **74**, 035101 (2006).
[8] A. Marini, C. Hogan, M. Grüning, and D. Varsano, Comp. Phys. Comm. **180**, 1392 (2009).
[9] C. Friedrich, S. Blügel, and A. Schindlmayr, Phys. Rev. B **81**, 125102 (2010).
[10] L. Martin-Samos and G. Bussi, Comp. Phys. Comm **180**, 1416 (2009).
[11] J. Wilhelm, M. Del Ben, and J. Hutter, J. Chem. Theory and Comput. **12**, 3623 (2016).
[12] F. Bruneval, T. Rangel, S. M. Hamed, M. Shao, C. Yang, and J. B. Neaton, Comput. Phys. Commun. **208**, 149 (2016).
[13] J. A. Berger, L. Reining, and F. Sottile, Phys. Rev. B **82**, 041103 (2010).
[14] R. W. Godby and R. J. Needs, Phys. Rev. Lett. **62**, 1169 (1989).
[15] W. von der Linden and P. Horsch, Phys. Rev. B **37**, 8351 (1988).
[16] G. E. Engel and B. Farid, Phys. Rev. B **47**, 15931 (1993).
[17] P. Liu, M. Kaltak, J. Klimeš, and G. Kresse, Phys. Rev. B **94**, 165109 (2016).
[18] V. Vlcek, E. Rabani, D. Neuhauser, and R. Baer, J. Chem. Theory Comput. **13**, 4997 (2017).
[19] P. Umari, G. Stenuit, and S. Baroni, Phys. Rev. B **79**, 201104 (2009).
[20] P. Umari, G. Stenuit, and S. Baroni, Phys. Rev. B **81**, 115104 (2010).
[21] A. Baldereschi and E. Tosatti, Solid State Commun. **29**, 131 (1979).

[22] R. Car, E. Tosatti, S. Baroni, and S. Leelaprute, Phys. Rev. B **24**, 985 (1981).

[23] D. Lu, F. Gygi, and G. Galli, Phys. Rev. Lett. **100**, 147601 (2008).

[24] F. Giustino, M. L. Cohen, and S. G. Louie, Phys. Rev. B **81**, 115105 (2010).

[25] J. Laflamme Janssen, B. Rousseau, and M. Côté, Phys. Rev. B **91**, 125120 (2015).

[26] H. F. Wilson, F. Gygi, and G. Galli, Phys. Rev. B **78**, 113303 (2008).

[27] H. F. Wilson, D. Lu, F. Gygi, and G. Galli, Phys. Rev. B **79**, 245106 (2009).

[28] H.-V. Nguyen, T. A. Pham, D. Rocca, and G. Galli, Phys. Rev. B **85**, 081101 (2012).

[29] T. A. Pham, H.-V. Nguyen, D. Rocca, and G. Galli, Phys. Rev. B **87**, 155148 (2013).

[30] M. Shao, L. Lin, C. Yang, F. Liu, F. H. Da Jornada, J. Deslippe, and S. G. Louie, Sci. China Math. **59**, 1593 (2016).

[31] M. Del Ben, F. H. da Jornada, A. Canning, N. Wichmann, K. Raman, R. Sasanka, C. Yang, S. G. Louie, and J. Deslippe, Comput. Phys. Commun. **235**, 187 (2019).

[32] S. L. Adler, Phys. Rev. **126**, 413 (1962).

[33] N. Wiser, Phys. Rev. **129**, 62 (1963).

[34] D. L. Johnson, Phys. Rev. B **9**, 4475 (1974).

[35] F. Gygi and A. Baldereschi, Phys. Rev. B **34**, 4405 (1986).

[36] P. Carrier, S. Rohra, and A. Görling, Phys. Rev. B **75**, 205126 (2007).

[37] J. Spencer and A. Alavi, Phys. Rev. B **77**, 193110 (2008).

[38] N. A. W. Holzwarth and X. Xu, Phys. Rev. B **84**, 113102 (2011).

[39] S. Lebègue, B. Arnaud, M. Alouani, and P. E. Bloechl, Phys. Rev. B **67**, 155208 (2003).

[40] R. W. Godby, M. Schlüter, and L. J. Sham, Phys. Rev. B **37**, 10159 (1988).

[41] M. Giantomassi, M. Stankovski, R. Shaltaf, M. Grüning, F. Bruneval, P. Rinke, and G.-M. Rignanese, Phys. Status Solidi B **248**, 275 (2011).

[42] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. D. Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, J. Phys.: Condens. Matter **21**, 395502 (2009).

[43] N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991).

[44] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevB.99.125128 for a detailed description of the computational parameters employed in the calculations.

[45] M. M. Rieger, L. Steinbeck, I. White, H. Rojas, and R. Godby, Comput. Phys. Commun. **117**, 211 (1999).

[46] A. Fleszar and W. Hanke, Phys. Rev. B **56**, 10228 (1997).

[47] J. Klimeš, M. Kaltak, and G. Kresse, Phys. Rev. B **90**, 075125 (2014).

[48] W. G. Aulbur, M. Städele, and A. Görling, Phys. Rev. B **62**, 7121 (2000).

[49] M. Rohlfing, P. Krüger, and J. Pollmann, Phys. Rev. B **48**, 17791 (1993).

[50] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[51] H. N. Rojas, R. W. Godby, and R. J. Needs, Phys. Rev. Lett. **74**, 1827 (1995).

[52] H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1976).

[53] K. Lejaeghere, G. Bihlmayer, T. Björkman, P. Blaha, S. Blügel, V. Blum, D. Caliste, I. E. Castelli, S. J. Clark, A. Dal Corso, S. de Gironcoli, T. Deutsch, J. K. Dewhurst, I. Di Marco, C. Draxl, M. Dułak, O. Eriksson, J. A. Flores-Livas, K. F. Garrity, L. Genovese, P. Giannozzi, M. Giantomassi, S. Goedecker, X. Gonze, O. Grånäs, E. K. U. Gross, A. Gulans, F. Gygi, D. R. Hamann, P. J. Hasnip, N. A. W. Holzwarth, D. Iuşan, D. B. Jochym, F. Jollet, D. Jones, G. Kresse, K. Koepernik, E. Küçükbenli, Y. O. Kvashnin, I. L. M. Locht, S. Lubeck, M. Marsman, N. Marzari, U. Nitzsche, L. Nordström, T. Ozaki, L. Paulatto, C. J. Pickard, W. Poelmans, M. I. J. Probert, K. Refson, M. Richter, G.-M. Rignanese, S. Saha, M. Scheffler, M. Schlipf, K. Schwarz, S. Sharma, F. Tavazza, P. Thunström, A. Tkatchenko, M. Torrent, D. Vanderbilt, M. J. van Setten, V. Van Speybroeck, J. M. Wills, J. R. Yates, G.-X. Zhang, and S. Cottenier, Science **351**, aad3000 (2016).

[54] D. R. Hamann, Phys. Rev. B **88**, 085117 (2013).

[55] A. Marini, G. Onida, and R. Del Sole, Phys. Rev. Lett. **88**, 016403 (2001).

[56] R. Courths and S. Hüfner, Phys. Rep. **112**, 53 (1984).

[57] J. I. Mustafa, M. Bernardi, J. B. Neaton, and S. G. Louie, Phys. Rev. B **94**, 155105 (2016).

[58] A. Grüneis, G. Kresse, Y. Hinuma, and F. Oba, Phys. Rev. Lett. **112**, 096401 (2014).

[59] T. Rangel, D. Kecik, P. E. Trevisanutto, G.-M. Rignanese, H. Van Swygenhoven, and V. Olevano, Phys. Rev. B **86**, 125125 (2012).

[60] S. Ismail-Beigi, Phys. Rev. B **73**, 233103 (2006).

[61] D. Y. Qiu, F. H. da Jornada, and S. G. Louie, Phys. Rev. Lett. **111**, 216805 (2013), Erratum, *ibid.* **115**, 119901 (2015).

[62] F. Hüser, T. Olsen, and K. S. Thygesen, Phys. Rev. B **88**, 245309 (2013).

[63] D. Y. Qiu, F. H. da Jornada, and S. G. Louie, Phys. Rev. B **93**, 235435 (2016).

[64] F. H. da Jornada, D. Y. Qiu, and S. G. Louie, Phys. Rev. B **95**, 035109 (2017).

[65] M. Schlipf and F. Gygi, Comput. Phys. Commun. **196**, 36 (2015).

[66] G. J. Martyna and M. E. Tuckerman, J. Chem. Phys. **110**, 2810 (1999).

[67] E. Maggio, P. Liu, M. J. van Setten, and G. Kresse, J. Chem. Theory Comput. **13**, 635 (2017).

[68] M. J. van Setten, F. Caruso, S. Sharifzadeh, X. Ren, M. Scheffler, F. Liu, J. Lischner, L. Lin, J. R. Deslippe, S. G. Louie, C. Yang, F. Weigend, J. B. Neaton, F. Evers, and P. Rinke, J. Chem. Theory Comput. **11**, 5665 (2015).

[69] B.-C. Shih, Y. Xue, P. Zhang, M. L. Cohen, and S. G. Louie, Phys. Rev. Lett. **105**, 146401 (2010).

[70] S. Tsoi, X. Lu, A. K. Ramdas, H. Alawadhi, M. Grimsditch, M. Cardona, and R. Lauck, Phys. Rev. B **74**, 165203 (2006).

[71] H. Alawadhi, S. Tsoi, X. Lu, A. K. Ramdas, M. Grimsditch, M. Cardona, and R. Lauck, Phys. Rev. B **75**, 205207 (2007).

[72] C. Friedrich, M. C. Müller, and S. Blügel, Phys. Rev. B **83**, 081101 (2011).

[73] C. Friedrich, M. C. Müller, and S. Blügel, Phys. Rev. B **84**, 039906(E) (2011).
[74] M. Stankovski, G. Antonius, D. Waroquiers, A. Miglio, H. Dixit, K. Sankaran, M. Giantomassi, X. Gonze, M. Côté, and G.-M. Rignanese, Phys. Rev. B **84**, 241201 (2011).
[75] A. Miglio, D. Waroquiers, G. Antonius, M. Giantomassi, M. Stankovski, M. Côté, X. Gonze, and G.-M. Rignanese, Eur. Phys. J. B **85**, 1 (2012).
[76] M. Zhang, S. Ono, N. Nagatsuka, and K. Ohno, Phys. Rev. B **93**, 155116 (2016).
[77] P. Larson, M. Dvorak, and Z. Wu, Phys. Rev. B **88**, 125205 (2013).
[78] G. Samsonidze, C.-H. Park, and B. Kozinsky, J. Phys. Condens. Matter **26**, 475501 (2014).
[79] J. A. Berger, L. Reining, and F. Sottile, Phys. Rev. B **85**, 085126 (2012).
[80] M. Usuda, N. Hamada, T. Kotani, and M. van Schilfgaarde, Phys. Rev. B **66**, 125101 (2002).
[81] L. Y. Lim, S. Lany, Y. J. Chang, E. Rotenberg, A. Zunger, and M. F. Toney, Phys. Rev. B **86**, 235113 (2012).
[82] B. D. Malone and M. L. Cohen, J. Phys. Cond. Matter **25**, 105503 (2013).
[83] M. J. van Setten, M. Giantomassi, X. Gonze, G.-M. Rignanese, and G. Hautier, Phys. Rev. B **96**, 155207 (2017).