All-electron self-consistent GW in the Matsubara-time domain: Implementation and benchmarks of semiconductors and insulators

Iek-Heng Chu,[1] Jonathan P. Trinastic,[1] Yun-Peng Wang,[1] Adolfo G. Eguiluz,[2] Anton Kozhevnikov,[3] Thomas C. Schulthess,[3] and Hai-Ping Cheng[1,*]

[1]*Department of Physics and Quantum Theory Project, University of Florida, Gainesville, Florida 32611, USA*
[2]*Department of Physics and Astronomy, The University of Tennessee, Knoxville, Tennessee 37996, USA*
[3]*Institute for Theoretical Physics, ETH Zurich, 8093 Zurich, Switzerland*

(Received 6 February 2016; published 28 March 2016)

The GW approximation is a well-known method to improve electronic structure predictions calculated within density functional theory. In this work, we have implemented a computationally efficient GW approach that calculates central properties within the Matsubara-time domain using the modified version of ELK, the full-potential linearized augmented plane wave (FP-LAPW) package. Continuous-pole expansion (CPE), a recently proposed analytic continuation method, has been incorporated and compared to the widely used Padé approximation. Full crystal symmetry has been employed for computational speedup. We have applied our approach to 18 well-studied semiconductors/insulators that cover a wide range of band gaps computed at the levels of single-shot G_0W_0, partially self-consistent GW_0, and fully self-consistent GW (full-GW), in conjunction with the diagonal approximation. Our calculations show that G_0W_0 leads to band gaps that agree well with experiment for the case of simple s-p electron systems, whereas full-GW is required for improving the band gaps in 3d electron systems. In addition, GW_0 almost always predicts larger band gap values compared to full-GW, likely due to the substantial underestimation of screening effects as well as the diagonal approximation. Both the CPE method and Padé approximation lead to similar band gaps for most systems except strontium titantate, suggesting that further investigation into the latter approximation is necessary for strongly correlated systems. Moreover, the calculated cation d band energies suggest that both full-GW and GW_0 lead to results in good agreement with experiment. Our computed band gaps serve as important benchmarks for the accuracy of the Matsubara-time GW approach.

DOI: 10.1103/PhysRevB.93.125210

I. INTRODUCTION

Calculations using density functional theory [1,2] (DFT) have become the standard *ab initio* technique to study the electronic and structural properties of molecules, nanoparticles, and periodic solids [3–6]. However, it is well known that the electronic band gap of semiconductors and insulators is severely underestimated within DFT due to the lack of a derivative discontinuity in standard exchange-correlation potentials [7]. This deficiency hinders the theory's useful application in fields such as optics, photovoltaics, thermoelectrics, and transport that require an accurate characterization of excited state properties.

The GW approximation, originally proposed by Hedin [8], provides a route to improve electronic descriptions and band gap results using many-body perturbation theory. The central quantity in this approach is the exchange-correlation self-energy (Σ^{xc}), which incorporates (i) the exact electronic exchange interaction and (ii) the complex electron-electron correlation accounting for screening effects often treated within the random phase approximation (RPA) [9,10]. This approach has been applied to a wide variety of materials and provides electronic structure results in better agreement with experiments compared to its DFT counterpart [11–17].

Although studies employing the GW approximation have enjoyed early success in improving band gap predictions, many implementations rely on the pseudopotential (PP) approximation that treats pseudo wave functions and valence-core interactions at the level of DFT [18–21]. To avoid the PP approximation, several all-electron GW implementations have been reported in recent years based on the full-potential linearized augmented plane wave (FP-LAPW) [22–26], the linearized muffin-tin orbital (LMTO) [26], and the projector-augmented wave (PAW) [27] in conjunction with a plane-wave basis [13,28]. Most of these all-electron studies have only implemented the G_0W_0 approximation due to its lower computational cost [26,29,30]; however, these single-shot calculations are plagued by violations of momentum, energy, and particle conservation laws [31–33]. They also introduce a troubling dependence on the choice of Kohn-Sham (K-S) basis used as a zeroth-order starting point [15,34]. Fully self-consistent GW, in which the single-particle Green's function (G) and self-energy (Σ, defined in the GW framework) are iterated by solving the Dyson equation to full self-consistency, avoids these issues and provide an unbiased physical picture predicted by GW theory. To date, few studies have performed self-consistent GW calculations within an all-electron framework [22,35,36], with one study performing the self-consistent GW method within the Matsubara-time domain [22,37,38], as first implemented by Ku and Eguiluz [22]. In that method, Σ and hence G have been approximated diagonally in the K-S basis, which is known as the diagonal approximation. However, this approach has only been applied to bulk Si and Ge, and its applicability to other semiconductors and insulators requires further examination.

There are two main advantages of performing GW calculations within the Matsubara-time domain. First, Σ is

*cheng@qtp.ufl.edu

simply the product of the single-particle Green's function (G) and screened Coulomb interaction (W). In contrast, the solution for Σ in Matsubara-frequency space requires a convolution of G and W that usually demands more frequency points to reach convergence [39]. Second, the Green's function in Matsubara-time lacks singular points that can arise in frequency space, which leads to smoother single-particle Green's functions compared to those in the frequency domain. Despite these advantages, the need for a reliable analytic continuation technique makes accurate calculations within Matsubara-time particularly challenging. The Padé approximation is often adopted for this purpose due to its simple implementation and low computational effort [40]. In this approach, the quantities of interest (e.g., Σ and G) are expressed as fractional polynomials that are fitted to computed values in the Matsubara-frequency domain. Such expressions are then analytically continued into the real-frequency domain. The reliability of this approximation remains under debate, and recently Staar and co-workers have proposed the continuous-pole expansion (CPE) as an alternative algorithm for analytic continuation from the Matsubara-frequency to the real-frequency domain [41]. Unlike the Padé approximation, this method explicitly takes into account the physical causality that places a constraint on the self-energy.

In this paper, we build upon an all-electron GW code we have already developed [11,42] by calculating Σ within the Matsubara-time domain, which improves the code's computational efficiency and provides self-consistent GW (full-GW) calculations within the diagonal approximation. We implement this method in conjunction with the CPE to solve for the quasiparticle energies in the real-frequency domain. We validate this method by investigating the electronic band gaps of a wide range of semiconductors and insulators at different levels of GW approximation. Our calculations demonstrate that the band gaps for $3d$ electron systems are often in better agreement with experiment when using full-GW than the commonly used G_0W_0 approximation, whereas the latter approximation often yields reasonable experimental agreement in simple s-p electron systems. We also find that both the CPE and Padé approximation yield very similar electronic band gaps among most tested systems; however, the CPE method provides a better electronic description of strongly correlated strontium titanate.

The rest of the paper is organized as follows. Section II outlines the full-GW approximation and Sec. III describes its implementation within the existing all-electron DFT package. Results and discussion are then presented in Sec. V, followed by the conclusion in Sec. VI.

II. BASICS OF THE THEORY

Within the single-particle picture, the excitation properties of solids can be determined by the single-particle Green's function via the Dyson equation. When expressed in the real-space and Matsubara-time domain, the Dyson equation reads

$$G(\mathbf{r},\mathbf{r}'|\tau) = G^0(\mathbf{r},\mathbf{r}'|\tau) + \int_0^\beta d\tau_1 \int_0^\beta d\tau_2 \int d\mathbf{r}_1 \int d\mathbf{r}_2 \times G^0(\mathbf{r},\mathbf{r}_1|\tau-\tau_1)\Delta\Sigma(\mathbf{r}_1,\mathbf{r}_2|\tau_1-\tau_2)G(\mathbf{r}_2,\mathbf{r}'|\tau_2), \quad (1)$$

where G and G^0 are the Green's functions associated with the interacting system of interest and a preselected reference system, respectively. In this work, the noninteracting K-S system calculated within DFT is adopted as the reference system. τ is the Matsubara-time argument that in general falls within $[-\beta,\beta]$ where $\beta = 1/k_BT$, k_B is Boltzmann's constant, and T is the temperature. Given that G obeys the relation $G(\mathbf{r},\mathbf{r}',-\tau) = -G(\mathbf{r},\mathbf{r}',-\tau+\beta)$ for $\tau \in [0,\beta]$, it is sufficient to restrict our study to $\tau \in [0,\beta]$. $\Delta\Sigma$ is the change in the electron-electron interaction between the interacting and reference K-S systems:

$$\Delta\Sigma(\mathbf{r},\mathbf{r}'|\tau) = \Sigma(\mathbf{r},\mathbf{r}'|\tau) - \Sigma^0(\mathbf{r},\mathbf{r}')\delta(\tau), \quad (2)$$

$$\Sigma(\mathbf{r},\mathbf{r}'|\tau) = \Sigma^H(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}')\delta(\tau) + \Sigma^{xc}(\mathbf{r},\mathbf{r}'|\tau), \quad (3)$$

$$\Sigma^H(\mathbf{r}) = \int d\mathbf{r}_1 \frac{\rho(\mathbf{r}_1)}{|\mathbf{r}-\mathbf{r}_1|}, \quad (4)$$

$$\rho(\mathbf{r}) = G(\mathbf{r},\mathbf{r}|\tau \to 0^-). \quad (5)$$

Here, Σ is the electron self-energy that captures the complicated electron-electron interactions. It is composed of the Hartree (Σ^H) and exchange-correlation (Σ^{xc}) components of the self-energy. Σ^H relates to the updated electronic charge density (ρ) and Σ^0 is the sum of Hartree and exchange-correlation potentials in the reference K-S system. $\delta(\tau)$ is the Dirac delta function.

Given the high computational cost of calculating Σ^{xc}, the standard method used to find this quantity is the GW approximation, which can be expressed in real-space and Matsubara-time as [37]

$$\Sigma^{xc}(\mathbf{r},\mathbf{r}'|\tau) = -G(\mathbf{r},\mathbf{r}'|\tau)\cdot W(\mathbf{r},\mathbf{r}'|\tau). \quad (6)$$

Here, W is the dynamically screened Coulomb potential, which describes the interactions between quasiparticles while including screening effects. The screened Coulomb potential obeys the Dyson equation that reads

$$W(\mathbf{r},\mathbf{r}'|\tau) = v(\mathbf{r},\mathbf{r}')\delta(\tau) + \int_0^\beta d\tau' \int d\mathbf{r}_1 \int d\mathbf{r}_2 v(\mathbf{r},\mathbf{r}_1) \times P(\mathbf{r}_1,\mathbf{r}_2|\tau-\tau')W(\mathbf{r}_2,\mathbf{r}'|\tau'), \quad (7)$$

where $v(\mathbf{r},\mathbf{r}') = 1/|\mathbf{r}-\mathbf{r}'|$ is the bare Coulomb potential, and P is the irreducible polarization within the RPA:

$$P(\mathbf{r},\mathbf{r}'|\tau) = G(\mathbf{r},\mathbf{r}'|\tau)\cdot G(\mathbf{r},\mathbf{r}'|-\tau). \quad (8)$$

In addition, $\Sigma^{xc}(\tau)$ is often expressed as the sum of the exchange self-energy, $\Sigma^x(\tau) = -G(\tau)\cdot v\delta(\tau)$, which corresponds to the Fock exchange term, and the correlation self-energy, $\Sigma^c = -G(\tau)\cdot[W(\tau) - v\delta(\tau)]$. Note that the self-energy in the Matsubara-time domain is simply a product of the Green's function and screened Coulomb potential, in contrast to the corresponding expression in the Matsubara-frequency domain that requires a convolution of G and W. The electron self-energy within the GW approximation [its exchange-correlation part is given in Eq. (6)] correlates with the Green's function, and thus both need to be solved self-consistently via Eq. (1).

The set of intercorrelated equations presented above allows us to compute G and Σ self-consistently. Once they are

converged to the required accuracy, a Fourier transform of Σ from the Matsubara-time to Matsubara-frequency domain is performed, i.e., $\{\Sigma(\tau)\} \rightarrow \{\Sigma(i\omega_n)\}$ where $\{\omega_n = (2n+1)\pi/\beta\}$ are the Matsubara frequencies with n being integers, and the spatial dependence of Σ is neglected for simplicity. This is then followed by an analytic continuation to real frequency space, $\{\Sigma(i\omega_n)\} \rightarrow \{\Sigma(\omega + i\eta)\}$, with η being a positive infinitesimal number, which yields the Green's function in the real-frequency domain and the excitation spectrum of the system.

III. IMPLEMENTATION OF THE SELF-CONSISTENT GW METHOD

In this section, we describe the full-GW approach in the Matsubara-time domain with the diagonal approximation, which has been implemented in the modified version of the ELK FP-LAPW package [42,43]. The approach is essentially similar to the one proposed by Ku and Eguiluz [22], but with more efficient computational schemes. In particular, (i) we have employed the more efficient uniform power mesh (UPM) in the Matsubara-time domain as proposed by Stan *et al.* [44], (ii) we have adopted the CPE for analytic continuation in conjunction with our full-GW method, and (iii) full crystal symmetry has been taken into account to significantly reduce the computational load. We briefly summarize these improvements in the subsections below.

A. Matsubara-time sampling

The Green's function G in the Matsubara-time domain varies smoothly in the range $0 \leqslant \tau \leqslant \beta$ and does not have any singularity points; however, it varies rapidly near $\tau = 0$ and β. To capture this behavior without losing computational efficiency, we employ the UPM to sample the τ axis on the grid $\{\tau_0 = 0, \tau_1, \tau_2, \ldots, \tau_M = \beta\}$ as proposed by Stan *et al.* [44], which is a modified version of the original one by Ku and Eguiluz [22]. The UPM grid can be characterized by a pair of integers (p, m) as well as the length of the interval β, in which p is the number of nonuniform subintervals generated between 0 and β with $2m-1$ evenly distributed grid points inside each of these sub-intervals. A UPM mesh with given (p, m) results in $2pm + 1$ grid points (including the endpoints) in the interval. In this scheme, the grid density increases for values of τ closer to the endpoints in order to capture the varying behavior of G. Using this scheme, explicit evaluation of quantities such as the self-energy and Green's function, which is normally computationally expensive, now only requires a coarse UPM grid. Thus, implementation of this grid significantly reduces the computational effort. For τ domain integrals that require knowledge of the integrand on a dense uniform τ grid, e.g., solving the Dyson equation, a higher-order interpolation such as cubic spline can be subsequently applied.

B. Scheme for full-GW in Matsubara-time domain

In this work, we expand and compute the Green's function G and self-energy Σ using the K-S basis ($\{\Psi_{n\mathbf{k}}(\mathbf{r})\}$ in spinor form, i.e. $\Psi_{n\mathbf{k}}(\mathbf{r}) = [\psi^{\uparrow}_{n\mathbf{k}}(\mathbf{r}), \psi^{\downarrow}_{n\mathbf{k}}(\mathbf{r})]$), whereas we evaluate the polarization function P and the screened Coulomb potential W in reciprocal space ($\{\mathbf{G}\}$). We also adopt the diagonal approximation such that Σ and G become approximately diagonal in the K-S basis, significantly reducing computational effort. This approximation has been shown to provide reasonable results for a variety of systems [13,26,29]. A direct generalization to include off-diagonal elements of Σ, i.e., removing the diagonal approximation, will be completed in the future. The full-GW approach is outlined below.

1. Green's function in the reference K-S system G^0

As a first step in full-GW, we construct the Green's function in the reference K-S system (G^0):

$$G^0_j(\mathbf{k}|\tau) = -\exp(-\epsilon_{j\mathbf{k}}\tau)[1 - n_F(\epsilon_{j\mathbf{k}})], \quad 0 \leqslant \tau \leqslant \beta \tag{9}$$

where $\{\epsilon_{j\mathbf{k}}\}$ are the K-S eigenenergies measured from the chemical potential μ of the system, $n_F = [\exp(\beta\epsilon_{j\mathbf{k}}) + 1]^{-1}$ is the Fermi-Dirac distribution, and $\mathbf{k}$ is a wave vector. In the zero-temperature limit, the results for a system with a nonzero band gap are insensitive to the choice of μ provided that it is placed inside the gap.

2. Irreducible polarization

The irreducible polarization P in the reciprocal space $\{\mathbf{G}\}$ can be obtained via Fourier transformations in Eq. (8) that reads,

$$P_{\mathbf{G}\mathbf{G}'}(\mathbf{q}|\tau) = \frac{1}{\Omega}\int d\mathbf{r}\int d\mathbf{r}' e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} P(\mathbf{r},\mathbf{r}'|\tau) e^{i(\mathbf{q}+\mathbf{G}')\cdot\mathbf{r}'}, \tag{10}$$

where Ω is the volume of the unit cell, and $\mathbf{q}$ falling within the first Brillouin zone (BZ). Using the relation between P and G in real space via Eq. (8), and by transforming the Green's function from the K-S basis to real space,

$$G(\mathbf{r},\mathbf{r}'|\tau) = \sum_{\mathbf{k}}^{\mathrm{BZ}} \sum_{j,\sigma} \psi^{\sigma}_{j\mathbf{k}}(\mathbf{r}) G_j(\mathbf{k}|\tau)[\psi^{\sigma}_{j\mathbf{k}}(\mathbf{r}')]^*, \tag{11}$$

it is straight-forward to show that the irreducible polarization in the reciprocal space can be expressed as follows,

$$P_{\mathbf{G}\mathbf{G}'}(\mathbf{q}|\tau) = \frac{1}{\Omega N_{\mathbf{k}}} \sum_{\mathbf{k}}^{\mathrm{BZ}} \sum_{j_1,j_2} M^{\mathbf{k}}_{j_2 j_1}(\mathbf{G},\mathbf{q}) \times Q_{j_1 j_2}(\mathbf{k},\mathbf{q}|\tau)\left[M^{\mathbf{k}}_{j_2 j_1}(\mathbf{G}',\mathbf{q})\right]^*, \tag{12}$$

$$\begin{aligned} Q_{j_1 j_2}(\mathbf{k},\mathbf{q}|\tau) &= G_{j_1}(\mathbf{k}+\mathbf{q}|\tau) G_{j_2}(\mathbf{k}|-\tau), \\ M^{\mathbf{k}}_{nm}(\mathbf{G},\mathbf{q}) &= \sum_{\sigma}\int d\mathbf{r}\left[\psi^{\sigma}_{n\mathbf{k}}(\mathbf{r})\right]^* e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} \psi^{\sigma}_{m\mathbf{k}+\mathbf{q}}(\mathbf{r}). \end{aligned} \tag{13}$$

Here, j_1 and j_2 are dummy band indices that run through both valence and conduction bands, σ is the dummy spin index, $\mathbf{q}$ is a reciprocal vector, and $\mathbf{G}$ is a reciprocal lattice vector. It is clear that the irreducible polarization P at any two distinct τ_1 and τ_2 in $[0, \beta]$ are decoupled. Therefore, parallelization over τ can be performed efficiently when P is evaluated.

3. Screened Coulomb potential

The screened Coulomb potential (W) can be computed once P is determined. Instead of directly solving for W,

during which the emergence of the Dirac delta function $\delta(\tau)$ (see Eq. (7)) may lead to numerical instability, we work with $\tilde{W}(\tau) \equiv W(\tau) - v\delta(\tau)$ (only τ dependence is indicated for simplicity). This formulation yields a correlation self-energy, $\Sigma^c(\tau) = -G(\tau)\cdot\tilde{W}(\tau)$, and exchange self-energy, $\Sigma^x(\tau) = -G(\tau)v\cdot\delta(\tau)$, such that $\Sigma^{xc}(\tau) = \Sigma^x(\tau) + \Sigma^c(\tau)$. In reciprocal space and Matsubara-time domain, $\tilde{W}$ obeys the following Dyson equation

$$\tilde{W}_{\mathbf{G}\mathbf{G}'}(\mathbf{q}|\tau) = \sum_{\mathbf{G}_2}\left[\sum_{\mathbf{G}_1} v_{\mathbf{G}\mathbf{G}_1}(\mathbf{q})P_{\mathbf{G}_1\mathbf{G}_2}(\mathbf{q}|\tau)\right]v_{\mathbf{G}_2\mathbf{G}'}(\mathbf{q}) + \int_0^\beta d\tau' \sum_{\mathbf{G}_2}\left[\sum_{\mathbf{G}_1} v_{\mathbf{G}\mathbf{G}_1}(\mathbf{q})P_{\mathbf{G}_1\mathbf{G}_2}(\mathbf{q}|\tau-\tau')\right] \times \tilde{W}_{\mathbf{G}_2\mathbf{G}'}(\mathbf{q}|\tau'), \tag{14}$$

where $v_{\mathbf{G}\mathbf{G}'}(\mathbf{q}) = 4\pi\delta_{\mathbf{G}\mathbf{G}'}/|\mathbf{q}+\mathbf{G}|^2$ is the Fourier transform of the bare Coulomb potential. We follow the algorithm proposed by Stan *et al.* [44] to discretize the τ-axis using the generated UPM grid. The above equation can then be re-arranged to form a linear matrix equation that reads

$$\sum_{r=0}^{M}\sum_{\mathbf{G}_2}\left[\delta_{\mathbf{G}\mathbf{G}_2}\delta_{p,r} - A_{\mathbf{G}\mathbf{G}_2}(\mathbf{q}|\tau^{(p)}-\tau^{(r)})\Delta\tau^{(r)}\right]\tilde{W}_{\mathbf{G}_2\mathbf{G}'}(\mathbf{q}|\tau^{(r)}) = \sum_{\mathbf{G}_2} A_{\mathbf{G}\mathbf{G}_2}(\mathbf{q}|\tau^{(p)})v_{\mathbf{G}_2\mathbf{G}'}(\mathbf{q}),$$

$$A_{\mathbf{G}\mathbf{G}_2}(\mathbf{q}|\tau) \equiv \sum_{\mathbf{G}_1} v_{\mathbf{G}\mathbf{G}_1}(\mathbf{q})P_{\mathbf{G}_1\mathbf{G}_2}(\mathbf{q}|\tau). \tag{15}$$

Here, the increments $\Delta\tau$ are positive, with $\Delta\tau^{(i)} = (\tau^{i+1} - \tau^{i-1})/2$ for $1 \leqslant i \leqslant M-1$. At the endpoints, $\Delta\tau^{(0)} = (\tau^1 - \tau^0)/2$ and $\Delta\tau^{(M)} = (\tau^M - \tau^{M-1})/2$.

4. Evaluating the self-energy

With $\tilde{W}(\tau)$ and $G(\tau)$ in hand, the correlation self-energy (Σ) can be evaluated as

$$\Sigma^c_n(\mathbf{k}|\tau) = -\frac{1}{\Omega N_\mathbf{k}}\sum_{\mathbf{q}}^{\mathrm{BZ}}\sum_{\mathbf{G}\mathbf{G}'}\sum_j \left[M^{\mathbf{k}-\mathbf{q}}_{jn}(\mathbf{G},\mathbf{q})\right]^* \times O^{\mathbf{G}\mathbf{G}'}_j(\mathbf{k},\mathbf{q}|\tau)M^{\mathbf{k}-\mathbf{q}}_{jn}(\mathbf{G}',\mathbf{q}),$$

$$O^{\mathbf{G}\mathbf{G}'}_j(\mathbf{k}-\mathbf{q}|\tau) = G_j(\mathbf{k}-\mathbf{q}|\tau)\tilde{W}(\mathbf{q}|\tau). \tag{16}$$

On the other hand, the exchange self-energy Σ^x is evaluated in real-space due to the slow convergence of Σ^x in reciprocal space [11],

$$\Sigma^x_{n\mathbf{k}} = -\sum_{\mathbf{k}'\in\mathrm{BZ}}\sum_m^{\mathrm{occ}}\int d\mathbf{r}\sum_\sigma \left[\psi^\sigma_{n\mathbf{k}}(\mathbf{r})\right]^*\psi^\sigma_{m\mathbf{k}'}(\mathbf{r}) \times \int d\mathbf{r}'\frac{\sum_{\sigma'}\left[\psi^{\sigma'}_{m\mathbf{k}'}(\mathbf{r}')\right]^*\psi^{\sigma'}_{n\mathbf{k}}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}f_{m\mathbf{k}'}, \tag{17}$$

where $f_{j\mathbf{k}} = G_j(\mathbf{k}|0^-)$ is the occupation number of the K-S eigenfunction $\Psi_{j\mathbf{k}}(\mathbf{r})$. Similarly, the Hartree potential is expressed as

$$\Sigma^H_{n\mathbf{k}} = \sum_\sigma\int d\mathbf{r}\left|\psi^\sigma_{n\mathbf{k}}(\mathbf{r})\right|^2 \times \int d\mathbf{r}'\frac{\sum_{\mathbf{k}'\in\mathrm{BZ}}\sum_{\sigma',m}\left|\psi^{\sigma'}_{m\mathbf{k}'}(\mathbf{r}')\right|^2}{|\mathbf{r}-\mathbf{r}'|}f_{m\mathbf{k}'}. \tag{18}$$

5. Dressed Green's function

During the full-GW calculation, the Green's function (G) is updated in each iteration using the newly obtained self-energy Σ in the Dyson equation, which reads

$$G^N_j(\mathbf{k}|\tau) = G^0_j(\mathbf{k}|\tau) + \int_0^\beta d\tau_2 Z_{j\mathbf{k}}(\tau,\tau_2)G^N_j(\mathbf{k}|\tau_2), \tag{19}$$

$$Z_{j\mathbf{k}}(\tau,\tau_2) = Z^x_{j\mathbf{k}}(\tau,\tau_2) + Z^c_{j\mathbf{k}}(\tau,\tau_2), \tag{20}$$

$$Z^x_{j\mathbf{k}}(\tau,\tau_2) = G^0_j(\mathbf{k}|\tau-\tau_2)\left[\Sigma^x_{N,j}(\mathbf{k}) + \Sigma^H_{N,j}(\mathbf{k}) - \Sigma_{0,j}(\mathbf{k})\right], \tag{21}$$

$$Z^c_{j\mathbf{k}}(\tau,\tau_2) = \int_0^\beta d\tau_1 G^0_j(\mathbf{k}|\tau-\tau_1)\cdot\Sigma^c_{N,j}(\mathbf{k}|\tau_1-\tau_2). \tag{22}$$

The integrals along the τ axis in Eqs. (19) and (22) may have substantial numerical errors when performed on the UPM mesh that becomes coarse farther away from the end points of $0 \leqslant \tau \leqslant \beta$. To overcome this issue, a cubic spline interpolation is applied to the Green's function and self-energy elements between two adjacent τ grid points, in which the increment $\Delta\tau$ in the resulting dense uniform τ grid is selected as $\tau_1 - \tau_0$. This is also the smallest $\Delta\tau$ in the UPM mesh. Then the Dyson equation is solved on the generated, denser uniform τ mesh. Similar to the algorithm for $\tilde{W}$ as proposed by Stan *et al.* [44], the Dyson equation for G along τ axis can be rearranged to form a linear matrix equation:

$$\sum_{r=1}^{N}[\delta_{p,r} - \Delta\tau^{(r)}Z_{j\mathbf{k}}(\tau^{(p)},\tau^{(r)})]G^N_j(\mathbf{k}|\tau^{(r)}) = G^0_j(\mathbf{k}|\tau^{(p)}). \tag{23}$$

During the full-GW calculation, we repeat the steps mentioned above in each iteration using the newly obtained Green's function G, as indicated in Eqs. (12), (15)–(18), and (21)–(23). We solve for the self-energy and the Green's function in the Matsubara-time domain self-consistently until any given accuracy is reached. Note that this corresponds to the single-shot G_0W_0 if the self-consistent calculation is terminated at the first iteration. The approximated calculation known as GW_0 can also be performed if the screened Coulomb potential W is kept constant after the first iteration, whereas G is updated during the self-consistent loop.

6. Analytic continuation

To obtain quantities that can be measured in experiments, such as the excitation spectrum, knowledge of G and Σ in the real-frequency domain is required. This is achieved by a two-step procedure performed after calculating the converged self-energy in the Matsubara-time domain [$\Sigma^{xc}(\tau)$]. First, a Fourier transformation from the Matsubara-time to

Matsubara-frequency domain is employed. For a given band index j and $\mathbf{k}$, this reads

$$\Sigma_j^{xc}(\mathbf{k}|i\omega_n) = \int_0^\beta d\tau\, e^{i\omega_n\tau}\Sigma_j^{xc}(\mathbf{k}|\tau), \tag{24}$$

where $\omega_n = (2n+1)\pi/\beta$ is the Matsubara frequency with n being an integer. We use a cubic spline interpolation of the UPM grid for the accurate evaluation of the integral. Second, we implement analytic continuation using the CPE method proposed by Staar *et al.* [41] to yield the self-energy in the real-frequency domain $[\Sigma^{xc}(\omega+i\eta)]$. Unlike the commonly used Padé approximation [40], where the self-energy elements are simply expanded as polynomials, the CPE takes advantage of the fact that the self-energy in the upper complex plane (z) can be expressed as

$$\Sigma_j^{xc}(\mathbf{k},z) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} d\omega \frac{\mathrm{Im}\left[\Sigma_j^{xc}(\mathbf{k},\omega+i\eta)\right]}{\omega - z}, \tag{25}$$

$$\mathrm{Im}\left[\Sigma_j^{xc}(\mathbf{k},\omega+i\eta)\right] < 0, \tag{26}$$

where η is a positive infinitesimal and Eq. (26) arises from causality. For each j and $\mathbf{k}$, $\mathrm{Im}\,\Sigma_j^{xc}(\mathbf{k}|\omega+i\eta)$ can be expanded as a set of piecewise linear functions of ω with undetermined coefficients $\{a_{nj}(\mathbf{k})\}$. This leads to $\Sigma_j^{xc}(\mathbf{k},z) = \sum_m a_{mj}(\mathbf{k})\Phi_{mj,\mathbf{k}}(z)$ where $\Phi_{mj,\mathbf{k}}(z)$ is some analytic function defined in the upper complex plane z. With the set of computed elements $\{\tilde{\Sigma}_j^{xc}(\mathbf{k}|i\omega_m)\}$ in hand and Eq. (26) as the constraint, for each given j and $\mathbf{k}$, $\{a_{nj}(\mathbf{k})\}$ are then determined by minimizing the norm function Ω defined as

$$\Omega_j(\mathbf{k}) = \sum_{m=0}^{M}\left|\tilde{\Sigma}_j^{xc}(\mathbf{k}|i\omega_m) - \Sigma_j^{xc}(\mathbf{k}|i\omega_m)\right|^2, \tag{27}$$

with M being the number of positive Matsubara frequencies. Given the fitted $\Sigma_j^{xc}(\mathbf{k}|\omega)$, for each j and $\mathbf{k}$, the Green's function associated with the interacting system can be determined using

$$G_j(\mathbf{k}|\omega) = \frac{1}{\left[G_j^0(\mathbf{k}|\omega)^{-1} - \Sigma_j^{xc}(\mathbf{k}|\omega) - \Sigma_j^H(\mathbf{k})\right]}. \tag{28}$$

The quasiparticle energies, and hence the electronic band gap, can be directly obtained from the spectral function $A_{j\mathbf{k}}(\omega) = -\frac{1}{\pi}\mathrm{Im}[G_{j\mathbf{k}}(\omega)]$ for given j and $\mathbf{k}$.

C. Use of crystal symmetry for computational speedup

Calculating the elements of $\Sigma^c(\tau)$ can be computationally expensive as it involves the evaluation of Eqs. (12), (13), (15), and (16). Such computational effort can be considerably reduced using crystal symmetry to decrease the number of required operations. The allowed crystal symmetry operations are those that leave the Hamiltonian invariant. Using these operations, reciprocal vectors in the first BZ $\{\mathbf{k}_{\mathrm{BZ}}\}$ are decomposed to a number of subsets. The reciprocal vectors in each of these subsets are related via the action of the symmetry operations. Therefore, the first BZ can be represented using a reduced set of $\mathbf{k}$ vectors that form the irreducible BZ (IBZ), denoted as $\{\mathbf{k}_{\mathrm{IBZ}}\}$.

Suppose $S_u \equiv \{(\mathbf{R}_i|\mathbf{t}_i), i = 1,\ldots,N_u\}$ is the set of symmetry operations in which $\mathbf{R}$ is a 3×3 rotation matrix and $\mathbf{t}$ the translation vector in real space. The application of a given symmetry operation, $B_i = (\mathbf{R}_i|\mathbf{t}_i)$, on the real-space vector $\mathbf{r}$ and reciprocal vector in IBZ lead respectively to

$$B_i\mathbf{r} = \mathbf{R}_i\mathbf{r} + \mathbf{t}_i, \tag{29}$$

$$\mathbf{k}_{\mathrm{BZ}} = B_i\mathbf{k}_{\mathrm{IBZ}} = \mathbf{R}_i\mathbf{k}_{\mathrm{IBZ}} + \mathbf{G}_{\mathbf{R}i}, \tag{30}$$

where $\mathbf{G}_{\mathbf{R}i}$ is the reciprocal lattice vector that brings $\mathbf{R}_i\mathbf{k}_{\mathrm{IBZ}}$ back to the first BZ. For a given $\mathbf{q}_{\mathrm{BZ}}$ that is associated with $\mathbf{q}_{\mathrm{IBZ}}$ via $\mathbf{R}$ and $\mathbf{G}$ using Eq. (30), it is straightforward to prove that the plane-wave matrix M in Eq. (13), the irreducible polarization $P(\tau)$ in Eq. (12), and $\tilde{W}(\tau)$ in Eq. (15) obey the following relations:

$$M_{nm}^{\mathbf{k}}(\mathbf{G},\mathbf{q}_{\mathrm{BZ}}) = M_{nm}^{\mathbf{R}^{-1}\mathbf{k}}[\mathbf{G}_1,\mathbf{q}_{\mathrm{IBZ}}] \times \exp[-i(\mathbf{R}\mathbf{q}_{\mathrm{IBZ}} + \mathbf{G}_{\mathbf{R}} + \mathbf{G})\cdot\mathbf{t}], \tag{31}$$

$$P_{\mathbf{G}\mathbf{G}'}(\mathbf{q}_{\mathrm{BZ}}|\tau) = P_{\mathbf{G}_1\mathbf{G}_1'}(\mathbf{q}_{\mathrm{IBZ}}|\tau)\exp[-i(\mathbf{G}-\mathbf{G}')\cdot\mathbf{t}], \tag{32}$$

$$\tilde{W}_{\mathbf{G}\mathbf{G}'}(\mathbf{q}_{\mathrm{BZ}}|\tau) = \tilde{W}_{\mathbf{G}_1\mathbf{G}_1'}(\mathbf{q}_{\mathrm{IBZ}}|\tau)\exp[-i(\mathbf{G}-\mathbf{G}')\cdot\mathbf{t}], \tag{33}$$

where $\mathbf{G}_1 = \mathbf{R}^{-1}(\mathbf{G}+\mathbf{G}_{\mathbf{R}})$ and $\mathbf{G}_1' = \mathbf{R}^{-1}(\mathbf{G}'+\mathbf{G}_{\mathbf{R}})$. It follows that the correlation self-energy can be rearranged as

$$\Sigma_n^c(\mathbf{k}|\tau) = -\frac{1}{\Omega N_{\mathbf{k}}}\sum_{\mathbf{q}_{\mathrm{IBZ}}}\sum_{\mathbf{R}}\sum_{\mathbf{G}\mathbf{G}'}\tilde{W}_{\mathbf{G}\mathbf{G}'}(\mathbf{q}_{\mathrm{IBZ}}|\tau) \times \sum_j\left[M_{jn}^{\mathbf{R}^{-1}\mathbf{k}-\mathbf{q}_{\mathrm{IBZ}}}(\mathbf{G},\mathbf{q}_{\mathrm{IBZ}})\right]^* \times G_j(\mathbf{R}^{-1}\mathbf{k}-\mathbf{q}_{\mathrm{IBZ}}|\tau)M_{jn}^{\mathbf{R}^{-1}\mathbf{k}-\mathbf{q}_{\mathrm{IBZ}}}(\mathbf{G}',\mathbf{q}_{\mathrm{IBZ}}). \tag{34}$$

Here, $\mathbf{R}^{-1}\mathbf{k} - \mathbf{q}_{\mathrm{IBZ}}$ is assumed to fall in the set of $\{\mathbf{k}_{\mathrm{BZ}}\}$ vectors. It is thus sufficient to compute the summands in the above equation for the sets of $\{\mathbf{q}_{\mathrm{IBZ}}\}$ and $\{\mathbf{k}_{\mathrm{BZ}}\}$ vectors, which leads to significant reduction to computational time. Similarly, the computation of the elements of Σ^x can be sped up with the use of symmetry operations for $\mathbf{k}$. According to Eq. (17), in particular, $\mathbf{k}$ associated with Σ^x can be confined to the IBZ, whereas $\mathbf{k}'$ runs over the first BZ.

IV. COMPUTATIONAL DETAILS

The full-GW scheme has been applied to calculate the electronic band gaps of 18 diverse semiconductors and insulators. We have adopted the experimental lattice parameters of 5.43 Å (Si), 5.658 Å (Ge), 5.66 Å (GaAs), 4.35 Å (SiC), 5.91 Å (CaSe), 3.57 Å (diamond), 5.64 Å (NaCl), 4.21 Å (MgO), 3.62 Å (cubic BN), 4.01 Å (LiF), 3.91 Å (cubic $SrTiO_3$), 4.27 Å (Cu_2O), 4.52 Å (GaN), 4.58 Å (zinc-blende ZnO), 5.42 Å (zinc-blende ZnS), 5.67 Å (zinc-blende ZnSe), 6.05 Å (zinc-blende CdSe), and 5.82 Å (zinc-blende CdS) throughout this work. All DFT calculations have been carried out using the modified version of the ELK FP-LAPW package [42,43]. The augmented plane wave + local orbitals (APW+lo) basis [45] with a single second-order local orbital per core or semicore state has been adopted. The local density approximation (LDA) [46] has been utilized for the exchange-correlation functionals. When expanding the interstitial potential and charge density, the maximum length of the reciprocal lattice

vector $|\mathbf{G}|$ has been chosen as 12 a.u. The angular momentum has been truncated as $\ell_{\max} = 8$ for the expansions of muffin-tin charge density, potential, and wave function. In the expansion of the wave function, $|\mathbf{G}+\mathbf{k}|_{\max} = 8.0/R_{\text{avg}}$ has been used, where R_{avg} is the average of the muffin-tin radii (R^{MT}) in each system. Linearization energy ($E_{\ell,\nu}$), which is associated with each radial function labeled with ν, is chosen at the center of the corresponding band with ℓ-like character. The first BZ has been sampled by a $4 \times 4 \times 4$ **k** mesh for all the systems except for diamond, where a $6 \times 6 \times 6$ **k** mesh has been used instead. All the aforementioned parameters have been carefully tested to achieve total energy convergence.

In the GW calculations, the cutoff for $|\mathbf{G}+\mathbf{q}|$ used in Eqs. (12) and (16) has been set 4.0 a.u. for all the systems except for the systems of ZnO, diamond, and cubic BN (c-BN), where a cutoff of 5.0 a.u. has been selected instead. These length cutoffs correspond to a kinetic-energy cutoff of 16 Ry and 25 Ry, respectively. The Matsubara-time (τ) domain has been sampled with a (9,5) UPM mesh, which consists of 91 grid points between 0 and β associated with an artificial temperature of 300 K. A minimum of 150 conduction bands have been included for the band summations in Eqs. (12) and (16) for the systems studied to ensure the convergence of the band gaps. In the GW_0 and full-GW calculations, states with an energy falling in the energy window of ± 15 eV around the DFT-LDA Fermi energy have been updated, and the number of iterations has been set to 4. In the transformation indicated in Eq. (24), a set of 128 positive Matsubara frequencies has been adopted, which is subsequently used in analytic continuation schemes of both the CPE and Padé approximation. For comparison, we have also performed G_0W_0 calculations using the plasmon-pole approximation (PPA), in which we have selected the model proposed by Godby and Needs [47] that has proven to be in consistent agreement with numerical integration method [11,48]. All the above parameters are carefully examined to ensure the band gap values converged to within 50 meV.

V. RESULTS AND DISCUSSION

A. Benchmarking Si, Ge, and GaAs

We first apply the Matsubara-time GW method to study the electronic properties of bulk silicon (Si), a prototypical system that has been studied as a benchmark for previous GW code developments. Figures 1(a) and 1(b) illustrate the Matsubara-time Green's functions $[G(\tau)]$ of the band-edge states at Γ_v and X_c at different levels of GW approximations, where K_v (K_c) denotes the highest occupied (lowest unoccupied) single-particle state at K. $G(\tau)$ approaches -1 and 0 at each end of the τ axis. For the case of the valence (conduction) band state in a semiconductor/insulator, $G(\beta^-) \to -1$ (0) to account for the occupation number of that state. It can be seen that, in the Matsubara-time domain, full-GW leads to substantial changes of G compared to G_0W_0. It is worth pointing out that the dressed G at Γ_v upon full-GW becomes very similar to that from the LDA, i.e., G^0. On the other hand, the full-GW leads to more deviation of G at X_c from G^0, suggesting that GW corrections to the conduction bands are likely more pronounced than to the valence bands. Figure 1(c) shows the typical Green's function in Matsubara-frequency domain (both real and imaginary parts) for the band edge states of bulk Si from full-GW calculations.

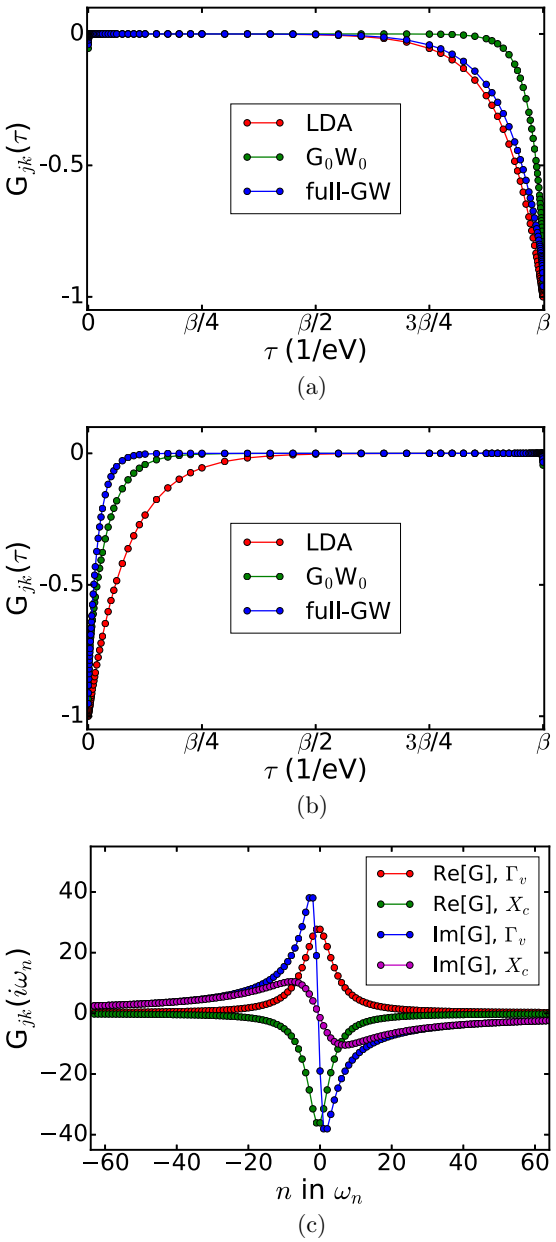

FIG. 1. Single-particle Green's functions $G_{n\mathbf{k}}$ at the (a) valence band maximum (Γ_v) and (b) conduction band minimum (X_c) of bulk Si in the Matsubara-time domain. (c) Single-particle Green's function of bulk Si in the Matsubara-frequency domain from full-GW calculations.

The calculated band gaps for bulk Si are tabulated in Table I. When the non-self-consistent G_0W_0 calculation is performed, the direct band gap at Γ and the indirect band gap from Γ to X are, respectively, 3.40 and 1.38 eV. These values are in relatively good agreement with experimental values [49]. Compared to those obtained from plane-wave pseudopotential (PP) based and/or all-electron G_0W_0, our computed G_0W_0 direct band gap and the indirect band gap are 0.1 and 0.2 eV higher, respectively. Moreover, we notice that band gap values

TABLE I. Band gap values of bulk Si for various levels of approximation. The values in parentheses are computed using the Padé approximation. All values are in eV.

	$\Gamma_v - \Gamma_c$	$\Gamma_v - X_c$
This work		
LDA	2.52	0.58
G_0W_0	3.40 (3.38)	1.38 (1.36)
GW_0	3.69 (3.68)	1.59 (1.58)
full-GW	3.41 (3.57)	1.44 (1.44)
Plane-wave PP, G_0W_0		
Tiago *et al.*[a]	3.24	1.18
All-electron, G_0W_0		
Hamada *et al.*[b]	3.30	1.14
Kotani *et al.*[c]	3.13	
Ku *et al.*[d]	3.12	
Gomez-Abal *et al.*[e]		1.15
All-electron, full-GW		
Ku *et al.*[d]	3.48	
Experiment[f]	3.35	1.25

[a]Reference [50].
[b]Reference [51].
[c]Reference [52].
[d]Reference [22].
[e]Reference [29].
[f]Reference [49].

are further increased by 0.2 eV upon implementing the partially self-consistent, GW_0 calculation. However, fully self-consistent GW brings the direct and indirect band gap values close to those calculated within the G_0W_0 approximation. Our full-GW results are also comparable with the previous study by Ku *et al.*, which uses a similar implementation to the present method. We also compare the results at different levels of GW using either the CPE method or Padé approximation. Band gap results using the Padé approximation generally agree well with those using CPE analytic continuation within 0.02 eV. However, the Si direct band gap value predicted by the Padé approximation is 0.16 eV higher than the CPE value, and is 0.09 eV higher than the value by Ku et al. This also shows that CPE results are generally in better agreement with experiment. In addition, all levels of GW calculations, from G_0W_0 to full-GW, overestimate the experimental indirect band gap value by 0.13 to 0.34 eV. The overestimation arising by full-GW also agrees with the previous GW study [13].

Note that the important effect of core electrons on the valence-core interaction, and hence exchange self-energy, has been discussed for bulk Si in the previous study [22]. We have also evaluated the exchange self-energy elements of band edge states Γ_v and X_c with and without the core electrons. The difference in self-energy can be as large as 2 eV, in line with values given in that study.

We have also compared the spectral functions $[A_{j\mathbf{k}}(\omega)]$ of the band edge states Γ_v and X_c of bulk Si from CPE to those obtained from the Padé approximation, as shown in Figs. 2(a) and 2(b) for the cases of G_0W_0 and full-GW, respectively. In the case of Si, results from these two approaches of analytic continuation are very similar in terms of peak position as well as the broadening of peaks that is related to the lifetime of the associated quasiparticle states.

Finally, we demonstrate the computational advantage of the current implementation by evaluating silicon's G_0W_0 band gaps with a similar parameter set but using a direct numerical

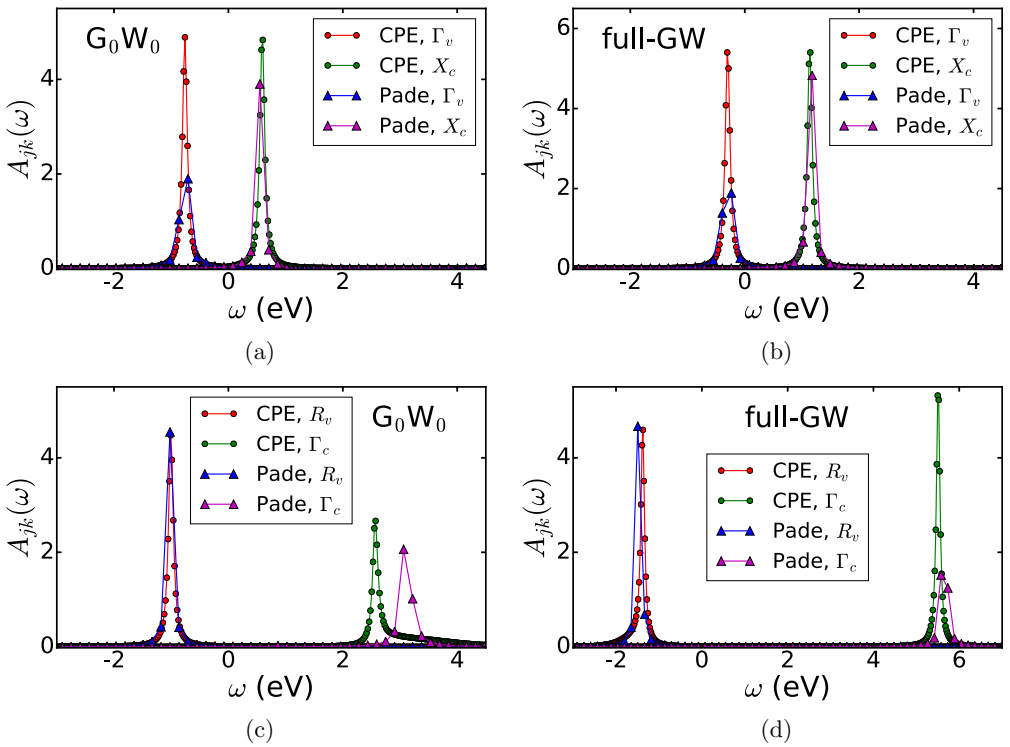

FIG. 2. Spectral functions of band edge states Γ_v and X_c of bulk Si from (a) G_0W_0 and (b) full-GW. Spectral functions of band edge states R_v and Γ_c of $SrTiO_3$ from (c) G_0W_0 and (d) full-GW.

TABLE II. Band gap values of bulk Ge for various levels of approximation. The values in parentheses are computed using the Padé approximation. All values are in eV.

	$\Gamma_v - \Gamma_c$	$\Gamma_v - L_c$	$\Gamma_v - X_c$
This work			
LDA	−0.19	0.03	0.64
G_0W_0	0.49 (0.51)	0.58 (0.59)	0.65 (0.70)
GW_0	1.09 (1.10)	0.85 (0.86)	1.35 (1.33)
full-GW	1.11 (1.11)	0.85 (0.85)	1.30 (1.30)
Plane-wave PP, G_0W_0			
Tiago *et al.*[a]	0.85	0.65	0.98
All-electron, G_0W_0			
Kotani *et al.*[b]	0.89	0.57	
Ku *et al.*[c]	1.11	0.51	0.49
All-electron, full-GW			
Ku *et al.*[c]	1.51	0.79	0.71
Experiment[d]	0.90	0.74	1.30

[a]Reference [50].
[b]Reference [52].
[c]Reference [22].
[d]Reference [49].

TABLE III. Band gap values of bulk GaAs for various levels of approximation. The values in parentheses are computed using the Padé approximation. All values are in eV.

	$\Gamma_v - \Gamma_c$	$\Gamma_v - L_c$	$\Gamma_v - X_c$
This work			
LDA	0.23	0.81	1.31
G_0W_0	1.48 (1.47)	1.62 (1.62)	1.98 (1.94)
GW_0	1.82 (1.83)	2.00 (2.00)	2.31 (2.30)
full-GW	1.80 (1.81)	1.95 (1.96)	2.23 (2.25)
Plane-wave PP, G_0W_0			
Tiago *et al.*[a]	1.38	1.65	1.83
All-electron, G_0W_0			
Kotani *et al.*[b]	1.20	1.40	1.46
Gomez-Abal *et al.*[c]	1.29		
Friedrich *et al.*[d]			
Experiment[e]	1.52	1.82	1.98

[a]Reference [50].
[b]Reference [52].
[c]Reference [29].
[d]Reference [24].
[e]Reference [53].

integration method in the real-frequency domain. We find that more than 1000 frequency points are needed to achieve the converged results. Since the computational load at each frequency or Matsubara-time grid point is similar, it is clear that significant computational speedup can be accomplished when GW calculation is performed in the Matsubara-time domain (91 τ points used in this work).

Table II summarizes the band gaps at different levels of theory for bulk Ge. The minimal, indirect band gap of bulk Ge is between Γ_v and L_c according to experiment [49]. It is clear that both LDA and G_0W_0 predict a minimal band gap as direct at Γ, inconsistent with experiment. It is only when the self-consistency is considered in GW (either GW_0 or full-GW) that the correct indirect band gap can be predicted. Note that the results from full-GW agree well with experimental data, and also very close to those from GW_0, regardless of the CPE or Padé approximation being adopted. It is worth pointing out that there is a substantial difference between our results and those by Ku *et al.*, with a band gap difference as large as 0.5 eV. We believe that such discrepancy is due mainly to the insufficient amount of empty bands used in their study, as pointed out in the previous study by Tiago *et al.* [50].

Gallium arsenide is another common compound we use as a benchmark, with computed band gap results shown in Table III. This compound has also been extensively investigated, which has a direct electronic band gap at Γ. Our calculations show that G_0W_0 results in the best agreement with experiment [53], and also agree with previous all-electron G_0W_0 studies with a $\sim$0.2 eV difference. Moreover, both full-GW and GW_0 lead to larger band gap values compared to the G_0W_0 results, and are overestimated by around 0.3 eV compared to experiment. Such trends regarding G_0W_0 and full-GW are also in line with previous GW studies within the plane-wave PAW potential framework [13]. Similar to the aforementioned compounds investigated, the CPE and Padé approximation lead to results that are very close to each other. Our full-GW results presented here also serve as important predictions for this level of theory since there are no previous all-electron-based, self-consistent GW results for GaAs.

In general, G_0W_0 accurately predicts Si and GaAs band gap values but predicts inaccurate bulk Ge band gap values compared to experiment. On the other hand, full-GW band gaps agree fairly well with experiment across all three elements, and GW_0 generally worsens the band gaps compared to full-GW.

B. Band gap calculations for other semiconductors and insulators

Having demonstrated the accuracy of full-GW calculations for predicting electronic band gaps in benchmark materials, we next report results for 18 semiconductors/insulators that have band gaps covering a wide range of values from less than 1 eV to over 10 eV. The calculated minimal band gaps are summarized in Table IV, comparing all levels of approximation, and also in Fig. 3, which visualizes LDA, G_0W_0, GW_0, and full-GW results. As expected, the LDA band gaps are always severely underestimated compared to experimental values. Upon GW corrections, the electronic band gaps for all the systems studied are substantially improved. In the following, we discuss the effects of G_0W_0 and full-GW band gap corrections by categorizing the compounds studied into three groups: (1) simple s-p electron systems involving Si, SiC, C, BN, LiF, NaCl, and MgO; (2) systems with deep d electron levels relative to the valence band maximum, including Ge, GaAs, GaN, CaSe, CdS, CdSe, ZnO, ZnS, and ZnSe; and (3) systems with relatively shallow d electron levels, including $SrTiO_3$ and Cu_2O.

Concerning simple s-p electron systems, the G_0W_0 corrected band gaps are in very good agreement with experimental data, with a relative band gap error of $\pm$10% for most compounds with the exception of diamond, for which G_0W_0

TABLE IV. Electronic band gap (in eV) of various semiconductors and insulators calculated by DFT-LDA, different levels of Matsubara-time GW (G_0W_0, GW_0, and full-GW), and PPA-G_0W_0. Values in the parentheses are obtained using the Padé approximation. The experimental values (Expt.) are also given for comparison.

	LDA	G_0W_0	GW_0	full-GW	PPA-G_0W_0	Expt.
Si	0.58	1.38 (1.36)	1.59 (1.58)	1.44 (1.44)	1.28	1.25[a]
Ge	0.03	0.58 (0.59)	0.85 (0.86)	0.85 (0.85)	0.71	0.74[a]
GaAs	0.24	1.48 (1.47)	1.82 (1.83)	1.80 (1.81)	1.51	1.52[b]
SiC	1.27	2.44 (2.45)	2.90 (2.90)	2.64 (2.56)	2.30	2.40[a]
CaSe	2.00	3.89 (3.94)	4.60 (4.64)	4.35 (4.34)	3.89	3.85[c]
C	4.14	6.15 (6.15)	6.42 (6.43)	6.10 (6.11)	6.09	5.48[a]
NaCl	4.74	8.09 (8.11)	9.00 (9.02)	8.27 (8.28)	8.11	8.5[d]
MgO	4.65	7.79 (7.78)	8.74 (8.74)	7.94 (7.94)	7.75	7.83[e]
BN	4.34	6.71 (6.73)	7.16 (7.18)	7.10 (7.11)	6.58	6.1–6.4[f]
LiF	8.94	14.51 (14.54)	15.78 (15.81)	14.45 (14.47)	14.55	14.20[g]
$SrTiO_3$	1.75	3.58 (4.08)	7.01 (7.13)	6.87 (7.22)	3.86	3.25[h]
Cu_2O	0.52	1.61 (1.54)	2.16 (2.17)	2.00 (2.02)	1.59	2.17[i]
GaN	1.70	3.01 (3.05)	3.61 (3.66)	3.36 (3.38)	3.03	3.27[j]
ZnO	0.60	2.31 (2.35)	3.69 (3.71)	3.53 (3.56)	2.32	3.44[k]
ZnS	1.80	3.46 (3.43)	4.06 (4.09)	3.92 (3.85)	3.43	3.91[k]
ZnSe	1.01	2.43 (2.48)	3.03 (3.09)	2.94 (2.96)	2.50	2.95[a]
CdS	0.86	2.01 (2.03)	2.63 (2.66)	2.49 (2.50)	2.06	2.50[a]
CdSe	0.34	1.42 (1.51)	1.97 (1.98)	1.92 (1.93)	1.46	1.83[a]

[a]Reference [49].
[b]Reference [53].
[c]Reference [54].
[d]Reference [55].
[e]Reference [56].
[f]Reference [57].
[g]Reference [58].
[h]Reference [59].
[i]Reference [60].
[j]Reference [61].
[k]Reference [62].

overestimates the experimental gap by 0.6 eV (12%). This may be attributed to the RPA that leads to more severe underestimation of the screening effect in diamond, as pointed out in a previous study [13]. Results using PPA are remarkably close to general G_0W_0 calculations, differing by only 0.1 eV or less. Our G_0W_0 band gaps are all comparable to previous all-electron G_0W_0 calculations [24,29]. When full self-consistency is taken into account, our calculations show that full-GW may further overestimate the electronic band gap due probably to the underestimated screening effect by RPA, in agreement with previous findings [13,65]. The exceptions are diamond and the ionic crystals NaCl and LiF, for which the inclusion of self-consistency tends to improve results. Furthermore, the band gaps at different levels of GW have also been computed based on the CPE and Padé approximation. According to our results, they are in remarkable agreement with each other, with a typical difference of 0.1 eV or less in all the cases. This also confirms the applicability of the Padé approximation and the analytic continuation approach for s-p electron systems.

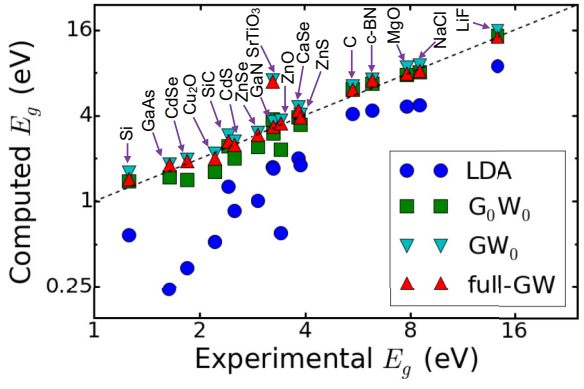

FIG. 3. Computed electronic band gap at DFT-LDA as well as GW levels versus the experimental counterpart for all the compounds studied in this work except for Ge. A logarithmic scale is adopted for both axes.

It is worth pointing out that the band gaps from partially self-consistent GW_0 are considerably higher than full-GW ones, which contrasts previous findings [13,66]. This different trend is likely due to differences in method implementation. Specifically, the Green's function under the diagonal approximation is fully updated during the GW_0 iteration in our approach. In contrast, previous GW_0 studies have only shifted the quasiparticle energies to update the presumedly diagonal Green's function [13], or within the Hermitian approximation to the full self-energy [66]. Such a different trend may also relate to the diagonal approximation for the

TABLE V. The average positions of cation outermost filled d bands at Γ (in eV) for GaAs, GaN, ZnO, ZnS, ZnSe, CdS, and CdSe with zinc-blende structure, as calculated using DFT-LDA, different levels of Matsubara-time GW (G_0W_0, GW_0, and full-GW) along with CPE. The experimental values (Expt.) are also given for comparison.

	LDA	G_0W_0	GW_0	full-GW	Expt.
GaAs	−14.91	−16.79	−19.37	−19.42	−(18.7–18.82)[a]
GaN	−13.62	−16.49	−18.54	−18.53	−17.0[b]
ZnO	−5.30	−7.21	−7.71	−7.64	−(7.5–8.81)[a]
ZnS	−6.30	−7.46	−9.31	−9.29	−9.03[a]
ZnSe	−6.55	−7.68	−9.59	−9.57	−(8.9–9.2)[a]
CdS	−7.62	−8.56	−9.71	−9.70	−(9.2–10.0)[a]
CdSe	−7.86	−8.70	−9.91	−9.91	−(9.9–10.7)[a]

[a]Reference [63].
[b]Reference [64].

Green's function in our approach. Compared to our full-GW results, the further overestimation of GW_0 band gaps may be attributed to the further underestimation of screening due to the screened potential (W), which is not updated iteratively within GW_0. This also highlights the importance of the full self-consistency.

For systems with deep 3d electrons, we have observed that G_0W_0 corrected band gaps are typically still underestimated. Fully self-consistent GW calculations are necessary to achieve better agreement with experimental data. The exception involves CaSe and GaAs, in which full-GW leads to overestimated band gaps by about 0.3–0.5 eV, corresponding to a relative difference of more than 12%. Compared to full-GW, GW_0 results in about 5% larger band gap values in the systems studied. The difference is smaller than in the case of s-p electron systems. Moreover, in the cases of ZnO and Cu_2O, it is clear that band gaps resulting from G_0W_0 are substantially underestimated by at least 0.5 eV compared to the experimental values. In particular, the G_0W_0 band gap of ZnO is 1 eV lower than the experimental data, agreeing well with previously underestimated values [13]. Upon full-GW, the band gaps of these systems are significantly improved such that they are within 0.2 eV of experimental results.

To study the systems with deep 3d electrons in more detail, we have computed the energies of the outermost cation filled d shell at Γ for GaAs, GaN, ZnO, ZnS, ZnSe, CdS, and CdSe (e.g., 4d and 3d for Cd and Zn, respectively). They are estimated as the average of all the corresponding d band energies, and are shown in Table V. For Cu_2O and $SrTiO_3$, given that the 3d bands are substantially hybridized and broadened, a simple averaging cannot give a clear picture so we will study them in more depth in a separate study. It is clear that the d band levels predicted by DFT-LDA are offset by at least 2 eV compared to the experimental data, as expected. On the other hand, G_0W_0 tends to improve the d band levels of all the systems studied, but the discrepancy can still be as large as 1 eV in systems such as GaAs and CdSe. Our G_0W_0 results are in good agreement with the previous all-electron study [30]. Upon self-consistency, we observe that full-GW leads to excellent agreement with experiment except for GaN, in which the deep Ga 3d band energy is substantially below the experimental value by 1.5 eV. It is worth pointing out that GW_0 leads to remarkably similar results to those from full-GW across all the systems.

Electronic structure predictions for the perovskite, $SrTiO_3$, provide our most inaccurate and intriguing results. Our G_0W_0 approach overestimates the band gap by about 0.3 eV, which is also consistent with the studies by Friedrich *et al.* [24] and Kang *et al.* [67], and both full-GW and GW_0 worsen the band gap prediction further with a result of more than 6.8 eV, much higher than the experimental value of 3.25 eV. We have further varied parameters such as the number of conduction bands and the cutoff of reciprocal lattice vectors, and the corresponding results only slightly change.

A previous GW study by Cappellini *et al.* also showed that the minimal band gap of $SrTiO_3$ can be severely overestimated even at the level of G_0W_0 (5.07 eV) [68], Such an overestimation may be attributed to the improper description of local field effects by their model dielectric function. Moreover, our full-GW band gap of $SrTiO_3$ is indeed in line with the previous findings, in which the band gap is overestimated by around 0.9 eV in all-electron quasiparticle self-consistent GW [69], whereas such overestimation becomes 1.8 eV in full-GW with the diagonal approximation in the plane-wave PAW potential framework [67]. Such a severe overestimation of the calculated full-GW band gap is thus likely due to the poor accuracy of the diagonal approximation adopted for G, which leads to unchanged charge and spin densities during full-GW. For systems with strongly correlated 3d electrons near the band edge, such as $SrTiO_3$, the quasiparticle wave functions may substantially deviate from K-S wave functions, resulting in considerable change in charge density and errors to the electronic band gap. Future work will include an investigation into how the diagonal approximation affects electronic structure predictions of transition metal oxides and other strongly correlated systems.

Another possibility is the missing electron-hole correlation effects in RPA [69]. Such effects have proven to be crucial in conjunction with self-consistency to predict correct electronic band gaps [65]. Further investigation excluding the diagonal approximation and/or including screening effects beyond RPA is necessary and will be conducted in the future. Similar to the other two types of systems, the CPE and Padé approximation lead to similar band gaps differing within 0.05 eV. The only exception is $SrTiO_3$, for which the band gap from both approaches can differ by as much as 0.5 eV, as indicated in Table IV and shown via the spectral functions in Figs. 2(c) and 2(d). Regarding the spectral functions of band edge states, difference in weight of spectral functions indicates that the estimated lifetime of the quasiparticle states may differ substantially. CPE appears to be the more valid method for analytic continuation given its general agreement with experiment for a wide range of systems. Still, the applicability of the Padé approximation is justified for many systems based on our calculations.

VI. CONCLUSION

To summarize, we have implemented an efficient Matsubara-time GW approach in conjunction with CPE, a newly developed analytic continuation method. The method has been used in a detailed study of the electronic band

gaps across 18 semiconductors and/or insulators at the levels of G_0W_0, GW_0, and full-GW approximations. Benchmark calculations of silicon's electronic structure demonstrate the accuracy and computational speedup of our Matsubara-time method compared to previously used frequency-domain calculations, indicating nearly an order of magnitude outperformance (in speed) of our Matsubara-time method over traditional frequency-domain GW approaches. Nevertheless, a systematic evaluation regarding the performance of our method is required and will be carried out in future work. Our results demonstrate that, for most of the simple s-p electron systems, G_0W_0 leads to reasonable agreement with experiments, and full-GW tends to overestimate the calculated band gaps, whereas full-GW is required for more accurate band gaps in the cases of 3d transition metal chalcogenides. These findings are in line with the previous GW studies, likely due to the underestimated screening effects by RPA during full-GW. We have also found that the band gap of strongly correlated systems such as $SrTiO_3$ can be substantially overestimated within the current framework, and off-diagonal elements in G as well as the electron-hole correlation effects beyond RPA may need to be included for more accurate results in those systems. Moreover, we have compared the results from both the CPE and Padé approximation. In general, CPE results are more consistently in agreement with experimental data in a wide range of systems, suggesting the applicability of CPE for analytic continuation as a standard for GW calculations. Finally, our calculations of average cation d band energies suggest that both full-GW and GW_0 lead to results in good agreement with experiment.

ACKNOWLEDGMENTS

This work is supported by the US Department of Energy (DOE) (USA), Office of Basic Energy Sciences (BES), under Contract No. DE-FG02-02ER45995. A.G.E. acknowledges funding support from National Science Foundation (NSF) (USA) under Grant No. OCI-0904972. All calculations were performed at the National Energy Research Scientific Computing Center (NERSC).

[1] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[2] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[3] W. Kohn, A. D. Becke, and R. G. Parr, J. Phys. Chem. **100**, 12974 (1996).
[4] R. M. Dreizler and E. K. Gross, *Density Functional Theory: An Approach to the Quantum Many-Body Problem* (Springer Science & Business Media, New York, 2012).
[5] J. K. Norskov, F. Abild-Pedersen, F. Studt, and T. Bligaard, Proc. Natl. Acad. Sci. **108**, 937 (2011).
[6] V. G. Ruiz, W. Liu, E. Zojer, M. Scheffler, and A. Tkatchenko, Phys. Rev. Lett. **108**, 146103 (2012).
[7] L. J. Sham and M. Schlüter, Phys. Rev. Lett. **51**, 1888 (1983).
[8] L. Hedin, Phys. Rev. **139**, A796 (1965).
[9] H. Eshuis, J. Bates, and F. Furche, Theor. Chem. Acc. **131**, 1084 (2012).
[10] X. Ren, P. Rinke, C. Joas, and M. Scheffler, J. Mater. Sci. **47**, 7447 (2012).
[11] I.-H. Chu, A. Kozhevnikov, T. C. Schulthess, and H.-P. Cheng, J. Chem. Phys. **141**, 044709 (2014).
[12] L. Yang, C.-H. Park, Y.-W. Son, M. L. Cohen, and S. G. Louie, Phys. Rev. Lett. **99**, 186801 (2007).
[13] M. Shishkin and G. Kresse, Phys. Rev. B **75**, 235102 (2007).
[14] G. Kresse, M. Marsman, L. E. Hintzsche, and E. Flage-Larsen, Phys. Rev. B **85**, 045205 (2012).
[15] F. Fuchs, J. Furthmüller, F. Bechstedt, M. Shishkin, and G. Kresse, Phys. Rev. B **76**, 115109 (2007).
[16] S. V. Faleev, M. van Schilfgaarde, and T. Kotani, Phys. Rev. Lett. **93**, 126406 (2004).
[17] T. Kotani, M. van Schilfgaarde, and S. V. Faleev, Phys. Rev. B **76**, 165106 (2007).
[18] J. Deslippe, G. Samsonidze, D. A. Strubbe, M. Jain, M. L. Cohen, and S. G. Louie, Comput. Phys. Commun. **183**, 1269 (2012).
[19] A. Marini, G. Onida, and R. Del Sole, Phys. Rev. Lett. **88**, 016403 (2001).
[20] H. Dixit, R. Saniz, D. Lamoen, and B. Partoens, Comput. Phys. Commun. **182**, 2029 (2011).
[21] T. A. Pham, H.-V. Nguyen, D. Rocca, and G. Galli, Phys. Rev. B **87**, 155148 (2013).
[22] W. Ku and A. G. Eguiluz, Phys. Rev. Lett. **89**, 126401 (2002).
[23] S. Sharma, J. K. Dewhurst, and C. Ambrosch-Draxl, Phys. Rev. Lett. **95**, 136402 (2005).
[24] C. Friedrich, S. Blügel, and A. Schindlmayr, Phys. Rev. B **81**, 125102 (2010).
[25] H. Jiang, R. I. Gmez-Abal, X.-Z. Li, C. Meisenbichler, C. Ambrosch-Draxl, and M. Scheffler, Comput. Phys. Commun. **184**, 348 (2013).
[26] M. Usuda, N. Hamada, T. Kotani, and M. van Schilfgaarde, Phys. Rev. B **66**, 125101 (2002).
[27] P. E. Blöchl, Phys. Rev. B **50**, 17953 (1994).
[28] M. Shishkin and G. Kresse, Phys. Rev. B **74**, 035101 (2006).
[29] R. Gómez-Abal, X. Li, M. Scheffler, and C. Ambrosch-Draxl, Phys. Rev. Lett. **101**, 106404 (2008).
[30] X.-Z. Li, R. Gmez-Abal, H. Jiang, C. Ambrosch-Draxl, and M. Scheffler, New J. Phys. **14**, 023006 (2012).
[31] G. Baym and L. P. Kadanoff, Phys. Rev. **124**, 287 (1961).
[32] G. Baym, Phys. Rev. **127**, 1391 (1962).
[33] N. E. Dahlen, R. van Leeuwen, and U. von Barth, Phys. Rev. A **73**, 012511 (2006).
[34] P. Rinke, A. Qteish, J. Neugebauer, C. Freysoldt, and M. Scheffler, New J. Phys. **7**, 126 (2005).
[35] F. Caruso, P. Rinke, X. Ren, M. Scheffler, and A. Rubio, Phys. Rev. B **86**, 081102 (2012).
[36] F. Caruso, P. Rinke, X. Ren, A. Rubio, and M. Scheffler, Phys. Rev. B **88**, 075105 (2013).
[37] G. D. Mahan, *Many-Particle Physics* (Springer Science & Business Media, New York, 2013).
[38] H. Bruus and K. Flensberg, *Many-Body Quantum Theory in Condensed Matter Physics: An Introduction* (Oxford University Press, Oxford, 2004).

[39] A. Kutepov, S. Y. Savrasov, and G. Kotliar, Phys. Rev. B **80**, 041103 (2009).
[40] H. Vidberg and J. Serene, J. Low Temp. Phys. **29**, 179 (1977).
[41] P. Staar, B. Ydens, A. Kozhevnikov, J.-P. Locquet, and T. Schulthess, Phys. Rev. B **89**, 245114 (2014).
[42] A. Kozhevnikov, A. G. Eguiluz, and T. C. Schulthess, in *SC'10, Proceedings of the 2010 ACM/IEEE International Conference for High Performance Computing, Networking, Storage, and Analysis* (IEEE Computer Society, Washington, DC, 2010), pp. 1–10.
[43] See https://code.google.com/p/exciting-plus for the all-electron DFT code. The complete GW code will soon be released at https://github.com/igator/.
[44] A. Stan, N. E. Dahlen, and R. van Leeuwen, J. Chem. Phys. **130**, 114105 (2009).
[45] E. Sjöstedt, L. Nordström, and D. Singh, Solid State Commun. **114**, 15 (2000).
[46] J. P. Perdew and Y. Wang, Phys. Rev. B **45**, 13244 (1992).
[47] R. W. Godby and R. J. Needs, Phys. Rev. Lett. **62**, 1169 (1989).
[48] 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).
[49] O. Madelung, M. Schulz, and H. Weiss, *Intrinsic Properties of Group IV Elements and IIIV, II-VI and I-VII Compounds*, Landolt-Börnstein, New Series, Vol. III (Springer-Verlag, New York, 1987).
[50] M. L. Tiago, S. Ismail-Beigi, and S. G. Louie, Phys. Rev. B **69**, 125212 (2004).
[51] N. Hamada, M. Hwang, and A. J. Freeman, Phys. Rev. B **41**, 3620 (1990).
[52] T. Kotani and M. van Schilfgaarde, Solid State Commun. **121**, 461 (2002).
[53] D. E. Aspnes, Phys. Rev. B **14**, 5331 (1976).
[54] Y. Kaneko and T. Koda, J. Cryst. Growth **86**, 72 (1988).
[55] R. T. Poole, J. G. Jenkin, J. Liesegang, and R. C. G. Leckey, Phys. Rev. B **11**, 5179 (1975).
[56] R. Whited, C. J. Flaten, and W. Walker, Solid State Commun. **13**, 1903 (1973).
[57] *Properties of Advanced Semiconductor Materials: GaN, AlN, InN, BN, SiC, and SiGe*, edited by M. E. Levinshtein, S. L. Rumyantsev, and M. S. Shur (Wiley, New York, 2001).
[58] M. Piacentini, D. W. Lynch, and C. G. Olson, Phys. Rev. B **13**, 5530 (1976).
[59] K. van Benthem, C. Elssser, and R. H. French, J. Appl. Phys. **90**, 6156 (2001).
[60] P. W. Baumeister, Phys. Rev. **121**, 359 (1961).
[61] H. Okumura, S. Yoshida, and T. Okahisa, Appl. Phys. Lett. **64**, 2997 (1994).
[62] C. Kittel and P. McEuen, *Introduction to Solid State Physics*, Vol. 8 (Wiley, New York, 1986).
[63] L. Ley, R. A. Pollak, F. R. McFeely, S. P. Kowalczyk, and D. A. Shirley, Phys. Rev. B **9**, 600 (1974).
[64] G. Martin, A. Botchkarev, A. Rockett, and H. Morkoç, Appl. Phys. Lett. **68**, 2541 (1996).
[65] M. Shishkin, M. Marsman, and G. Kresse, Phys. Rev. Lett. **99**, 246403 (2007).
[66] C. Rodl, F. Sottile, and L. Reining, Phys. Rev. B **91**, 045102 (2015).
[67] G. Kang, Y. Kang, and S. Han, Phys. Rev. B **91**, 155141 (2015).
[68] G. Cappellini, S. Bouette-Russo, B. Amadon, C. Noguera, and F. Finocchi, J. Phys. Condens. Matter **12**, 3671 (2000).
[69] M. van Schilfgaarde, T. Kotani, and S. Faleev, Phys. Rev. Lett. **96**, 226402 (2006).