All-electron periodic G_0W_0 implementation with numerical atomic orbital basis functions: Algorithm and benchmarks

Xinguo Ren [1,*] Florian Merz,[2] Hong Jiang,[3] Yi Yao [4,5] Markus Rampp [6] Hermann Lederer [6]
Volker Blum,[7,5] and Matthias Scheffler[8,9]

[1]*Beijing National Laboratory for Condensed Matter Physics, Institute of Physics, Chinese Academy of Sciences, 3rd South Str. 8, Beijing 100190, China*

[2]*Lenovo HPC Innovation Center, Meitnerstr. 9, D-70563 Stuttgart, Germany*

[3]*Beijing National Laboratory for Molecular Sciences, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China*

[4]*Department of Chemistry, University of North Carolina, Chapel Hill, North Carolina 27599, USA*

[5]*Department of Mechanical Engineering and Materials Science, Duke University, Durham, North Carolina 27708, USA*

[6]*Max Planck Computing and Data Facility, Giessenbachstrasse 2, D-85748 Garching, Germany*

[7]*Department of Chemistry, Duke University, Durham, North Carolina 27708, USA*

[8]*Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, D-14195 Berlin, Germany*

[9]*Physics Department, Humboldt-Universität zu Berlin, Zum Großen Windkanal 6, D-12489 Berlin, Germany*

We present an all-electron, periodic G_0W_0 implementation within the numerical atomic orbital (NAO) basis framework. A localized variant of the resolution-of-the-identity (RI) approximation is employed to significantly reduce the computational cost of evaluating and storing the two-electron Coulomb repulsion integrals. We demonstrate that the error arising from localized RI approximation can be reduced to an insignificant level by enhancing the set of auxiliary basis functions, used to expand the products of two single-particle NAOs. An efficient algorithm is introduced to deal with the Coulomb singularity in the Brillouin zone sampling that is suitable for the NAO framework. We perform systematic convergence tests and identify a set of computational parameters, which can serve as the default choice for most practical purposes. Benchmark calculations are carried out for a set of prototypical semiconductors and insulators, and compared to independent reference values obtained from an independent G_0W_0 implementation based on linearized augmented plane waves (LAPWs) plus high-energy localized orbitals (HLOs) basis set, as well as experimental results. With a moderate (FHI-aims *tier* 2) NAO basis set, our G_0W_0 calculations produce band gaps that typically lie in between the standard LAPW and the LAPW + HLO results. Complementing *tier* 2 with highly localized Slater-type orbitals (STOs), we find that the obtained band gaps show an overall convergence towards the LAPW + HLO results. The algorithms and techniques developed in this work pave the way for efficient implementations of correlated methods within the NAO framework.

I. INTRODUCTION

The electronic band structure determines the behavior of electrons and consequently a large variety of properties of periodic materials. Therefore, accurate computations of electronic band structures are crucial for *ab initio* quantum mechanical descriptions of materials. Kohn-Sham (KS) density-functional theory (DFT) [1,2] within its local-density [3,4] and generalized gradient approximations (LDA/GGAs) [5,6] offers a relatively inexpensive approach to compute band structures in solids, and has significantly improved our understanding of solid materials. However, LDA and GGAs do not include the correct underlying physics to describe experimentally relevant band structures, i.e., electron addition or removal energies and other quasiparticle properties. The substantial underestimation of the band gaps of insulating materials, and occasionally the incorrect description of energy level orderings have been a major drawback of this approach. Hybrid density functionals [7], especially those incorporating a portion of screened exact exchange [8], have gained increased popularity in condensed matter and materials science community, due to their overall improved band structure description. However, a certain level of empiricism is often required to tune the mixing and screening parameters to suitable values, and such functionals still lack the desired predictive power. Recently there has been progress to determine these semi-empirical parameters automatically in a self-adaptive way [9–13].

An alternative, and formally more rigorous approach to calculate electronic band structures is the Green-function-based many-body perturbation theory [14,15], whereby the band structure can be determined from the poles of the interacting Green function G of the system. Via the Dyson equation, G is linked to a reference, noninteracting Green function G_0 in terms of a non-Hermitian, dynamic self-energy, which

encompasses all the many-body exchange-correlation effects. In practice, approximations have to be made to calculate the self-energy, and a popular choice has been the so-called *GW* approximation [16–25], whereby the self-energy is given by a product of the Green function G and the screened Coulomb interaction W. Despite its simplicity, the *GW* approximation has been shown to yield significantly improved band structures, compared to their DFT counterparts, for a wide range of materials. Because the *GW* band structure is derived from a dynamical self-energy, and describes correctly the quasiparticle nature of single-particle excitations in materials, the *GW* band structure is often called *quasiparticle* band structure in the literature.

Since the 1980's, the *GW* approach has been implemented within different numerical frameworks, ranging from the pseudopotential plane wave method [18,26–28], to the projector augmented wave (PAW) method [29–33], a mixed representation of plane waves and real-space grid [34,35], the linearized muffin-tin orbital (LMTO) method [36], the linearized augmented plane wave (LAPW) method [37–41], and Gaussian orbital basis sets [42–44]. Needless to say, each implementation has its own advantages as well as limitations, but the considerable efforts behind these various implementations are clear indications of the importance and wide-spread influence of the *GW* method. As a matter of fact, the *GW* functionality is becoming a standard component of many widely used first-principles computational software packages.

In addition to the significant efforts devoted to its numerical implementations, the *GW* methodology itself is also under active development. In practice, *GW* calculations are often done on top of a preceding KS-DFT LDA/GGA calculation or a hybrid functional calculation within the generalized KS (gKS) scheme. In this case, the *GW* band structure is obtained as a one-shot correction to the KS-DFT one, and such a computational scheme is termed G_0W_0 in the literature. The G_0W_0 scheme frequently performs very well, but the obtained results obviously depend on the starting point, i.e., the approximation used in the preceding (generalized) KS-DFT calculation. Various schemes beyond the simple G_0W_0 approach have been proposed and tested, including, e.g., Green-function-only partially self-consistent *GW* (denoted as GW_0) [45,46], quasiparticle self-consistent *GW* (QPscGW) [47,48], and fully self-consistent *GW* [37,49–54], but G_0W_0 remains to be the most widely used approach in practical calculations.

The *GW* approach has been the state-of-the-art electronic-structure method for determining quasiparticle energies in semiconductors and insulators in the last three decades. In recent years, *GW* has become also popular as an approach to determine ionization energies and electron affinities of molecules [50,51,55–65]. Aside from its considerable importance in realistic description of organic materials, applying *GW* to molecules allows one to systematically benchmark the accuracy across different, independent *GW* implementations, as was done in the *GW* 100 test set project [64], and to benchmark the accuracy of the *GW* method against the traditional quantum chemistry approaches such as the coupled cluster (CC) method [60,63]. Most recently, an in-depth diagrammatic analysis has been carried out to compare and contrast *GW* with the equation-of-motion CC method [66].

Our own *GW* implementation [59] has been carried out within the all-electron numerical atomic orbital (NAO)-based Fritz Haber Institute *ab-initio molecular simulation* (FHI-aims) code package [67–71]. Comprehensive benchmark tests [72–74] indicate that FHI-aims offers impressive precision for ground-state DFT calculations. The *GW* implementation in FHI-aims was initially done for the G_0W_0 scheme and finite systems for nonperiodic boundary condition [59] but soon extended to the fully self-consistent *GW* [51,52] scheme. Like other correlated methods available in FHI-aims, our *GW* implementation is based on a technique known as variational density fitting [75,76] or the resolution-of-the-identity (RI) approximation [77–79], which expands the products of molecular orbitals (MOs) in terms of a set of auxiliary basis functions (ABFs). Our RI-based *GW* implementation, together with our special on-the-fly procedure for constructing the ABFs, turns out to be remarkably accurate, as demonstrated in the *GW* 100 project [64].

In this work, we extend our molecular G_0W_0 implementation to periodic systems. However, to extend a molecular *GW* implementation to periodic systems within the NAO framework, several numerical obstacles have to be overcome. Within the family of local orbitals, we are aware of only a few periodic *GW* implementations based on the LMTO method [36] or on Gaussian-type orbitals (GTOs) with the pseudopotential treatment of core ions [42–44], but to our knowledge there has been no reported NAO-based all-electron periodic *GW* implementation. In the case of ground-state KS-DFT, the numerical techniques for periodic implementations within the NAO framework have been developed over the years and is now well established, as can be seen by the availability of a number of NAO-based DFT codes [67,80–85]. When coming to correlated methods, which require unoccupied states and two-electron integrals, NAO-based implementations for periodic systems are still in their infancy. In the latter case, the major difficulties lie in the computation and storage of a large number of two-electron Coulomb repulsion integrals. Further challenges include how to construct efficient and high-quality NAO basis sets to describe the unoccupied energy states and how to treat the so-called Coulomb singularity in the Brillouin zone (BZ) sampling. In this paper, we will describe the algorithms and numerical techniques used in our periodic G_0W_0 implementation, focusing on our strategies to deal with the aforementioned challenging issues. Systematic convergence tests with respect to the computational parameters, and benchmarks against independent G_0W_0 implementations and experimental values for prototypical three-dimensional (3D) semiconductors and insulators will be reported as well. Although not discussed here, we would like to mention that our G_0W_0 implementation can be readily applied to one-dimensional (1D) and two-dimensional (2D) systems, because our basis functions are strictly localized in space and as detailed in Sec. III B, the Coulomb operator is truncated at the boundary of the supercell under the the Born–von Kármán (BvK) periodic boundary condition. Furthermore, our implementation is highly parallel, scaling up to tens of thousands of CPU cores. In this paper, we focus on the basic algorithms and numerical precision aspects, validating our implementation for 3D insulating systems. A discussion of the scalability and efficiency of implementation, as well as its performance for

low-dimensional systems, will be presented in a forthcoming paper.

This paper is organized as follows. In Sec. II we present the basic theory and algorithms behind our implementation, especially the working equations for periodic GW based on a localized variant of the RI approximation. In Sec. III the computational details will be discussed, including our procedure to generate the one-electron orbital basis functions and auxiliary basis functions, and our algorithm to treat the Coulomb singularity. Systematic convergence tests with respect to computational parameters are then presented in Sec. IV for Si and MgO. This is followed by Sec. V where benchmark calculations for a set of prototypical semiconductors and insulators are performed, and compared to well-established reference values. Finally we conclude our paper with an outlook in Sec. VI.

II. THEORETICAL FRAMEWORK

A. *GW* equations in real space

In this section, we recapitulate the basic equations of the commonly used G_0W_0 approach, where both the Green function G_0 and screened Coulomb interaction W_0 are determined by the orbitals and orbital energies, obtained from a preceding (g)KS-DFT calculation. In real space, the G_0W_0 self-energy is given by

$$\Sigma_\sigma^{G_0W_0}(\mathbf{r}, \mathbf{r}', i\omega)$$
$$= -\frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega' G_{0,\sigma}(\mathbf{r}, \mathbf{r}', i\omega - i\omega') W_0(\mathbf{r}, \mathbf{r}', i\omega'), \quad (1)$$

where the noninteracting Green function G_0 is

$$G_{0,\sigma}(\mathbf{r}, \mathbf{r}', i\omega) = \sum_{n,\mathbf{k}} w_{\mathbf{k}} \frac{\psi_{n,\sigma}^{\mathbf{k}}(\mathbf{r}) \psi_{n,\sigma}^{\mathbf{k}*}(\mathbf{r}')}{i\omega + \mu - \epsilon_{n,\sigma}^{\mathbf{k}}}. \quad (2)$$

Here $\psi_{n,\sigma}^{\mathbf{k}}$ and $\epsilon_{n,\sigma}^{\mathbf{k}}$ are KS orbitals and orbital energies, with n, σ being the orbital and spin indices, and $\mathbf{k}$ being a Bloch momentum vector in the first BZ (1BZ). Furthermore, $w_{\mathbf{k}}$ is integration weight of the $\mathbf{k}$ point (for even-spaced $\mathbf{k}$ grids, $w_{\mathbf{k}} = 1/N_{\mathbf{k}}$ with $N_{\mathbf{k}}$ being the number of $\mathbf{k}$ points in the 1BZ), ω is a frequency point on the imaginary frequency axis, and μ is the electronic chemical potential. The screened Coulomb interaction W_0 in Eq. (1) is defined as

$$W_0(\mathbf{r}, \mathbf{r}', i\omega) = \int d\mathbf{r}'' \varepsilon^{-1}(\mathbf{r}, \mathbf{r}'', i\omega) v(\mathbf{r}'', \mathbf{r}'), \quad (3)$$

where ε^{-1} is the inverse of the microscopic dielectric function ε. Within the random phase approximation (RPA), ε is fully determined by the independent particle response function χ_0 and the bare Coulomb interaction v,

$$\varepsilon(\mathbf{r}, \mathbf{r}', i\omega) = \delta(\mathbf{r} - \mathbf{r}') - \int d\mathbf{r}'' v(\mathbf{r}, \mathbf{r}'') \chi_0(\mathbf{r}'', \mathbf{r}', i\omega). \quad (4)$$

The noninteracting response function χ_0 can be expressed explicitly in terms of KS orbitals and orbital energies, according to the Adler-Wiser formula [86,87]

$$\chi_0(\mathbf{r}, \mathbf{r}', i\omega) = \sum_{m,n,\sigma} \sum_{\mathbf{k},\mathbf{q}}^{1BZ} w_{\mathbf{k}} w_{\mathbf{q}} \frac{\left(f_{m,\sigma}^{\mathbf{k+q}} - f_{n,\sigma}^{\mathbf{q}}\right) \psi_{m,\sigma}^{\mathbf{k+q}*}(\mathbf{r}) \psi_{n,\sigma}^{\mathbf{k}}(\mathbf{r}) \psi_{n,\sigma}^{\mathbf{k}*}(\mathbf{r}') \psi_{m,\sigma}^{\mathbf{k+q}}(\mathbf{r}')}{\epsilon_{m,\sigma}^{\mathbf{k+q}} - \epsilon_{n,\sigma}^{\mathbf{k}} - i\omega}. \quad (5)$$

In Eq. (5) the summations of the Bloch vectors $\mathbf{k}$, $\mathbf{q}$ are over the 1BZ, and $f_{n,\sigma}^{\mathbf{k}}$ are the orbitals' occupation factors. In our formulation and practical implementation to be described below, we work on the imaginary frequency axis. To get quasiparticle excitation energies, an analytical continuation of the self-energy from the imaginary to the real frequency axis will be carried out [34]. Alternatively, the G_0W_0 self-energy on the real frequency axis can also be directly computed via the contour deformation (CD) approach. For molecules and clusters this has already been implemented in FHI-aims [88]. It will be extended to the periodic case in the future.

B. Auxiliary basis representation of *GW* equations

In practical calculations, $\chi_0(\mathbf{r}, \mathbf{r}', i\omega)$ (as well as ε and W) has to be discretized either on a real-space grid [34] or more often it is expanded in terms of a suitable basis set. In the latter case, the basic choice to expand nonlocal quantities (χ_0, ε, and W_0) depends on the preceding computational framework to obtain the single-particle KS orbitals $\psi_{n,\sigma}^{\mathbf{k}}$. For example, in the pseudopotential plane-wave [18] or PAW frameworks [31,33], these basis functions are simply plane waves, and in the LMTO [89] or LAPW [38,40,41,90] frameworks, the so-called mixed product basis is used. Our own implementation employs the NAO basis function framework, whereby a set of atom-centered auxiliary basis functions is constructed to expand the products of two KS orbitals. In the molecular case, this can be expressed as

$$\psi_{m,\sigma}^*(\mathbf{r}) \psi_{n,\sigma}(\mathbf{r}) \approx \sum_\mu C_{m,n,\sigma}^\mu P_\mu(\mathbf{r}), \quad (6)$$

where $\psi_{n,\sigma}(\mathbf{r})$ is a molecular KS orbital, $P_\mu(\mathbf{r})$ is the μth auxiliary basis function, and $C_{m,n,\sigma}^\mu$ is the three-orbital (triple) expansion coefficient. It follows that χ_0, ε, and W can all be represented in terms of $P_\mu(\mathbf{r})$'s in a matrix form. These $P_\mu(\mathbf{r})$ basis functions are termed ABFs, which are distinct from the orbital basis sets (OBSs) $\{\varphi_i\}$ to expand a single KS orbital,

$$\psi_{n,\sigma}(\mathbf{r}) = \sum_i c_{i,n,\sigma} \varphi_i(\mathbf{r} - \boldsymbol{\tau}_i). \quad (7)$$

Here $c_{i,n,\sigma}$ are the KS eigenvectors, and φ_i is an orbital basis function (i.e., the NAO basis function mentioned above) centered at the atomic position $\boldsymbol{\tau}_i$. Throughout this paper we use m, n indices for denoting KS orbitals, i, j, k, l for atomic basis functions, and μ, ν, α, β for ABFs.

Our ABFs are also atom-centered, and they are constructed in a way similar to the mixed product basis in the LAPW framework, but without the plane-wave component in the interstitial region. In Ref. [59] we have described how to make use of the expansion in Eq. (6) to achieve efficient implementations of Hartree-Fock, second-order Møller-Plesset perturbation theory (MP2), the RPA, and *GW* within the NAO basis set framework. Approximations based on Eq. (6) are known as density fitting [75,76], or RI [77–79] in the context of evaluating two-electron Coulomb repulsion integrals, as already mentioned above. The RI-based implementation reported in Ref. [59], though quite accurate for both Hartree-Fock and correlated methods, is restricted to molecular geometries under nonperiodic boundary conditions. More recently, we have extended our formalism and implementation to periodic systems. A periodic, linear-scaling Hartree-Fock and screened exact exchange implementation was reported in Ref. [70]. In this paper, we focus on the extension of our *GW* implementation to periodic systems, based on the local variant of the RI approximation used in Refs. [69] and [70].

In periodic systems, the KS orbitals carry an additional **k** vector in their indices, and their products can be expanded in terms of the Bloch summation of the atom-centered ABFs,

$$\psi_{m,\sigma}^{\mathbf{k+q}*}(\mathbf{r})\psi_{n,\sigma}^{\mathbf{k}}(\mathbf{r}) = \sum_{\mu}^{N_{\text{aux}}} C_{m,n,\sigma}^{\mu}(\mathbf{k+q},\mathbf{k})P_{\mu}^{\mathbf{q}*}(\mathbf{r}),\quad (8)$$

where N_{aux} is the number of ABFs within each unit cell,

$$P_{\mu}^{\mathbf{q}}(\mathbf{r}) = \sum_{\mathbf{R}} P_{\mu}(\mathbf{r}-\mathbf{R}-\boldsymbol{\tau}_{\mu})e^{i\mathbf{q}\cdot\mathbf{R}},\quad (9)$$

and $C_{m,n,\sigma}^{\mu}(\mathbf{k+q},\mathbf{k})$ are the expansion coefficients which now depend on two independent Bloch wave vectors. In Eq. (9), $\boldsymbol{\tau}_{\mu}$ is the position of atom from which the μth ABF originates within the unit cell, and the sum runs over all unit cells **R** in the BvK supercell. In our implementation, the atom-centered ABFs are chosen to be real-valued, and hence

$P_{\mu}^{\mathbf{q}*}(\mathbf{r}) = P_{\mu}^{-\mathbf{q}}(\mathbf{r})$. Using Eqs. (5) and (8), one immediately arrives at

$$\chi_0(\mathbf{r},\mathbf{r}',i\omega) \approx \sum_{\mu,\nu}\sum_{\mathbf{q}} w_{\mathbf{k}} P_{\mu}^{\mathbf{q}*}(\mathbf{r})\chi_{0,\mu\nu}(\mathbf{q},i\omega)P_{\nu}^{\mathbf{q}}(\mathbf{r}'),\quad (10)$$

where the matrix representation of χ_0 is given by

$$\chi_{0,\mu\nu}(\mathbf{q},i\omega) = \sum_{\sigma,m,n}\sum_{\mathbf{k}} w_{\mathbf{k}}\frac{C_{m,n,\sigma}^{\mu}(\mathbf{k+q},\mathbf{k})C_{n,m,\sigma}^{\nu}(\mathbf{k},\mathbf{k+q})}{\epsilon_{m,\sigma}^{\mathbf{k+q}}-\epsilon_{n,\sigma}^{\mathbf{k}}-i\omega}.$$

$$(11)$$

To obtain the matrix representation of ε and W_0 in terms of ABFs, one still needs to compute the Coulomb matrix given by expanding the Coulomb operator in terms of the same set of ABFs,

$$V_{\mu\nu}(\mathbf{q}) = \int d\mathbf{r}d\mathbf{r}'\frac{P_{\mu}^{\mathbf{q}*}(\mathbf{r})P_{\nu}^{\mathbf{q}}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}.\quad (12)$$

The matrix form of ε and W_0 can then be obtained via matrix multiplication and inversion at each $(\mathbf{k},i\omega)$ point. For computational convenience, we use the symmetrized dielectric function $\tilde{\varepsilon} = v^{-1/2}\varepsilon v^{1/2}$, whose matrix form can be computed as

$$\tilde{\varepsilon}_{\mu\nu}(\mathbf{q},i\omega) = \delta_{\mu,\nu} - \sum_{\alpha\beta} V_{\mu\alpha}^{1/2}(\mathbf{q})\chi_{0,\alpha\beta}(\mathbf{q},i\omega)V_{\beta\nu}^{1/2}(\mathbf{q}),\quad (13)$$

where $V^{1/2}$ is the square root of the V matrix. The $\tilde{\varepsilon}$ matrix is then inverted and the matrix form of W_0 can be obtained as

$$W_{0,\mu\nu}(\mathbf{q},i\omega) = \sum_{\alpha,\beta} V_{\mu\alpha}^{1/2}(\mathbf{q})\tilde{\varepsilon}_{\alpha\beta}^{-1}(\mathbf{q},i\omega)V_{\beta\nu}^{1/2}(\mathbf{q}).\quad (14)$$

Noting that

$$W_{0,\mu\nu}(\mathbf{q},i\omega) = \iint d\mathbf{r}\,d\mathbf{r}' P_{\mu}^{\mathbf{q}*}(\mathbf{r})W_0(\mathbf{r},\mathbf{r}',i\omega)P_{\nu}^{\mathbf{q}}(\mathbf{r}')\quad (15)$$

and using Eqs. (1), (2), and (8), one arrives at the following expression for computing the diagonal matrix element of the G_0W_0 self-energy:

$$\Sigma_{n,\sigma}^{G_0W_0}(\mathbf{k},i\omega) = \iint d\mathbf{r}\,d\mathbf{r}' \psi_{n,\sigma}^{\mathbf{k}*}(\mathbf{r})\Sigma_{\sigma}^{G_0W_0}(\mathbf{r},\mathbf{r}',i\omega)\psi_{n,\sigma}^{\mathbf{k}}(\mathbf{r}') \quad (16)$$

$$= -\frac{1}{2\pi}\sum_{m,\mathbf{q}}\sum_{\mu,\nu}\int_{-\infty}^{\infty}d\omega'\frac{C_{n,m,\sigma}^{\mu}(\mathbf{k},\mathbf{k-q})W_{0,\mu\nu}(\mathbf{q},i\omega')C_{m,n,\sigma}^{\nu}(\mathbf{k-q},\mathbf{k})}{i\omega - i\omega' + \mu - \epsilon_{m,\sigma}^{\mathbf{k-q}}}. \quad (17)$$

In this formulation, the equations are well defined and the quantities can in principle be evaluated straightforwardly except at the Γ ($\mathbf{q} = 0$) point where elements of the V and W matrices between two nodeless s functions or between one nodeless s and one nodeless p function will diverge. Simply avoiding the Γ point in the BZ sampling is not an optimal solution since this can result in a prohibitively slow convergence with respect to the summation over the **k** points. Because of the localized, nonorthogonal nature of our ABFs, the Γ-point correction schemes developed in the context of plane-wave [91,92] or LAPW basis set framework [38] are not directly applicable here. We will discuss our treatment of the Γ-point singularity in Sec. III B.

C. Localized resolution of identity technique to determine the expansion coefficients

Obviously, in the above formulation, the key issues are (1) to construct a proper, sufficiently accurate auxiliary basis set $\{P_{\mu}^{\mathbf{q}}(\mathbf{r})\}$ and (2) to determine the expansion coefficients $C_{m,n,\sigma}^{\mu}(\mathbf{k+q},\mathbf{k})$ which are needed in the computation of both χ_0 matrix in Eq. (11) and G_0W_0 self-energy element in Eq. (17). Our procedure to construct the ABFs and the precision that can be achieved in practical calculations have been discussed in previous works [59,69,70]. We will come to this point only briefly in Sec. III A, in the context of periodic G_0W_0 calculations. In this subsection, we will

focus on point (2) and discuss how the expansion coefficients $C_{m,n,\sigma}^{\mu}(\mathbf{k} + \mathbf{q}, \mathbf{k})$ are determined.

In FHI-aims [67] the KS orbitals are expanded in terms of NAO basis functions,

$$\psi_{n\sigma}^{\mathbf{k}}(\mathbf{r}) = \sum_{i=1}^{N_\mathrm{b}} c_{i,n,\sigma}(\mathbf{k}) \varphi_i^{\mathbf{k}}(\mathbf{r})$$

$$= \sum_{i=1}^{N_\mathrm{b}} c_{i,n,\sigma}(\mathbf{k}) \sum_{\mathbf{R}} \varphi_i(\mathbf{r} - \mathbf{R} - \boldsymbol{\tau}_i) e^{i\mathbf{k}\cdot\mathbf{R}}, \qquad (18)$$

where $c_{i,n,\sigma}(\mathbf{k})$ are the KS eigenvectors, N_b is the number of basis functions within one unit cell, $\mathbf{R}$ is a Bravais lattice vector, and $\boldsymbol{\tau}_i$ is the position of the atom (within the unit cell) on which the basis function i is centered. In Eq. (18), the atomic function φ_i is given by

$$\varphi_i(\mathbf{r}) = u_{a,s,l}(r) Y_{lm}(\hat{\mathbf{r}}), \qquad (19)$$

where $u_{a,s,l}(r)$ is the radial function, and $Y_{lm}(\hat{\mathbf{r}})$ is a spherical harmonic [in our implementation real-valued harmonic, i.e., the real (cosine, positive m values) or imaginary (sine, negative m values) component of a spherical harmonic]. Thus the atomic orbital i is fully specified by the atom index a, the radial function index s, and the angular momentum indices l, m.

The ABFs used in FHI-aims are also atom-centered, but with different radial functions,

$$P_\mu(\mathbf{r}) = \xi_{a,s,l}(r) Y_{lm}(\hat{\mathbf{r}}), \qquad (20)$$

where $\xi_{a,s,l}(r)$ is the radial auxiliary function, and μ is also a combined basis index of a, s, l, and m. The radial auxiliary functions are constructed in a way [59,69] that the products of the KS orbitals $\psi_{m\sigma}\psi_{n\sigma}$, or equivalently the products of

orbital basis functions $\varphi_i\varphi_j$ can be represented by a linear combination of $\{P_\mu(\mathbf{r})\}$.

To compute the expansion coefficients $C_{m,n,\sigma}^{\mu}(\mathbf{k} + \mathbf{q}, \mathbf{k})$, defined in Eq. (8), we first determine the expansion coefficients of the products of two (Bloch summed) basis functions $\varphi_i^{\mathbf{k}+\mathbf{q}*}(\mathbf{r})\varphi_j^{\mathbf{k}}(\mathbf{r})$ in terms of the ABFs,

$$\varphi_i^{\mathbf{k}+\mathbf{q}*}(\mathbf{r})\varphi_j^{\mathbf{k}}(\mathbf{r}) = \sum_{\mu=1}^{N_\mathrm{aux}} \tilde{C}_{i,j}^{\mu}(\mathbf{k} + \mathbf{q}, \mathbf{k}) P_\mu^{\mathbf{q}*}(\mathbf{r}), \qquad (21)$$

and then transform $\tilde{C}_{i,j}^{\mu}(\mathbf{k} + \mathbf{q}, \mathbf{k})$ to the $C_{m,n,\sigma}^{\mu}(\mathbf{k} + \mathbf{q}, \mathbf{k})$ by multiplying with the KS eigenvectors,

$$C_{m,n,\sigma}^{\mu}(\mathbf{k} + \mathbf{q}, \mathbf{k}) = \sum_{i,j} c_{i,m,\sigma}^{*}(\mathbf{k} + \mathbf{q}) c_{j,n,\sigma}(\mathbf{k}) \tilde{C}_{i,j,\sigma}^{\mu}(\mathbf{k} + \mathbf{q}, \mathbf{k}).$$

$$(22)$$

Below we shall refer to $C_{m,n,\sigma}^{\mu}(\mathbf{k} + \mathbf{q}, \mathbf{k})$ as molecular orbital (MO) triple coefficients and $\tilde{C}_{i,j,\sigma}^{\mu}(\mathbf{k} + \mathbf{q}, \mathbf{k})$ as atomic orbital (AO) triple coefficients. Both types of triple coefficients depend on three orbital indices and in addition on two independent momentum vectors $\mathbf{k}$ and $\mathbf{q}$. Take the AO triple coefficients, for example: the number of entries scales as $N_\mathrm{b}(N_\mathrm{b} + 1)/2 N_\mathrm{aux} N_{\mathbf{k}}^2$, and it is quite expensive to compute and store them. In FHI-aims we can adopt the LRI approximation [69] to deal with this issue. Within the LRI approximation, the ABFs used to expand the product of two NAOs are restricted to those centering on the two atoms on which these two NAOs are centered. In quantum chemistry, such a two-center LRI scheme is also known as pair-atom RI (PARI) approximation [93,94]. In real space, the two NAOs i, j in general can originate from two different unit cells, labeled by two Bravais lattice vectors $\mathbf{R}_i$ and $\mathbf{R}_j$. The LRI approximation for periodic systems then implies that

$$\varphi_i(\mathbf{r} - \mathbf{R}_i - \boldsymbol{\tau}_i)\varphi_j(\mathbf{r} - \mathbf{R}_j - \boldsymbol{\tau}_j) \approx \sum_{\mu\in I} \tilde{C}_{i(\mathbf{R}_i),j(\mathbf{R}_j)}^{\mu(\mathbf{R}_i)} P_\mu(\mathbf{r} - \mathbf{R}_i - \boldsymbol{\tau}_i) + \sum_{\mu\in J} \tilde{C}_{i(\mathbf{R}_i),j(\mathbf{R}_j)}^{\mu(\mathbf{R}_j)} P_\mu(\mathbf{r} - \mathbf{R}_j - \boldsymbol{\tau}_j), \qquad (23)$$

where $\tilde{C}_{i(\mathbf{R}_i),j(\mathbf{R}_j)}^{\mu(\mathbf{R}_i)}$ are the two-center expansion coefficients where the lattice vector in parentheses associated with the basis index indicates the unit cell from which the basis function originates. Furthermore, I, J in Eq. (23) denote the atoms where φ_i and φ_j are centered, and $\mu \in I$ means the summation over the ABFs is restricted to those centering at the atom I. Because of the translational symmetry of the periodic system, one has $C_{i(\mathbf{R}_i),j(\mathbf{R}_j)}^{\mu(\mathbf{R}_i)} = C_{i(\mathbf{0}),j(\mathbf{R}_j-\mathbf{R}_i)}^{\mu(\mathbf{0})}$, where $\mathbf{0}$ here denotes the unit cell at the origin. Therefore Eq. (23) becomes

$$\varphi_i(\mathbf{r} - \mathbf{R}_i - \boldsymbol{\tau}_i)\varphi_j(\mathbf{r} - \mathbf{R}_j - \boldsymbol{\tau}_j) \approx \sum_{\mu\in I} \tilde{C}_{i(\mathbf{0}),j(\mathbf{R}_j-\mathbf{R}_i)}^{\mu(\mathbf{0})} P_\mu(\mathbf{r} - \mathbf{R}_i - \boldsymbol{\tau}_i) + \sum_{\mu\in J} \tilde{C}_{i(\mathbf{R}_i-\mathbf{R}_j),j(\mathbf{0})}^{\mu(\mathbf{0})} P_\mu(\mathbf{r} - \mathbf{R}_j - \boldsymbol{\tau}_j). \qquad (24)$$

This means that the two-center expansion coefficients in real space naturally split into two sectors, and each of them depends on only one independent lattice vector. Now, by Fourier transforming Eq. (24) to $\mathbf{k}$ space from both sides, we obtain

$$\varphi_i^{\mathbf{k}+\mathbf{q}*}(\mathbf{r})\varphi_j^{\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{R}_i,\mathbf{R}_j} e^{-i(\mathbf{k}+\mathbf{q})\cdot\mathbf{R}_i} e^{i\mathbf{k}\cdot\mathbf{R}_j} \varphi_i(\mathbf{r} - \mathbf{R}_i - \boldsymbol{\tau}_i)\varphi_j(\mathbf{r} - \mathbf{R}_j - \boldsymbol{\tau}_j)$$

$$\approx \sum_{\mathbf{R}_i,\mathbf{R}_j} e^{-i(\mathbf{k}+\mathbf{q})\cdot\mathbf{R}_i} e^{i\mathbf{k}\cdot\mathbf{R}_j} \left[\sum_{\mu\in I} \tilde{C}_{i(\mathbf{0}),j(\mathbf{R}_j-\mathbf{R}_i)}^{\mu(\mathbf{0})} P_\mu(\mathbf{r} - \mathbf{R}_i - \boldsymbol{\tau}_i) + \sum_{\mu\in J} \tilde{C}_{i(\mathbf{R}_i-\mathbf{R}_j),j(\mathbf{0})}^{\mu(\mathbf{0})} P_\mu(\mathbf{r} - \mathbf{R}_j - \boldsymbol{\tau}_j) \right]$$

$$= \sum_{\mu\in I} \left[\sum_{\mathbf{R}_i} e^{-i\mathbf{q}\cdot\mathbf{R}_i} P_\mu(\mathbf{r} - \mathbf{R}_i - \boldsymbol{\tau}_i) \sum_{\mathbf{R}_j} e^{i\mathbf{k}\cdot(\mathbf{R}_j-\mathbf{R}_i)} \tilde{C}_{i(\mathbf{0}),j(\mathbf{R}_j-\mathbf{R}_i)}^{\mu(\mathbf{0})} \right]$$

$$+ \sum_{\mu \in J} \left[\sum_{\mathbf{R}_j} e^{-i\mathbf{q}\cdot\mathbf{R}_j} P_\mu(\mathbf{r} - \mathbf{R}_j - \boldsymbol{\tau}_j) \sum_{\mathbf{R}_i} e^{-i(\mathbf{k}+\mathbf{q})\cdot(\mathbf{R}_i-\mathbf{R}_j)} \tilde{C}^{\mu(\mathbf{0})}_{i(\mathbf{R}_i-\mathbf{R}_j),j(\mathbf{0})} \right]$$

$$= \sum_{\mu \in I} \tilde{C}^{\mu(\mathbf{0})}_{i(-\mathbf{k}-\mathbf{q}),j(\mathbf{0})} P_\mu^{\mathbf{q}*}(\mathbf{r}) + \sum_{\mu \in J} \tilde{C}^{\mu(\mathbf{0})}_{i(\mathbf{0}),j(\mathbf{k})} P_\mu^{\mathbf{q}*}(\mathbf{r}). \tag{25}$$

Here we have introduced the notation

$$\tilde{C}^{\mu(\mathbf{0})}_{i(\mathbf{k}),j(\mathbf{0})} = \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} \tilde{C}^{\mu(\mathbf{0})}_{i(\mathbf{R}),j(\mathbf{0})}, \tag{26}$$

$$\tilde{C}^{\mu(\mathbf{0})}_{i(\mathbf{0}),j(\mathbf{k})} = \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} \tilde{C}^{\mu(\mathbf{0})}_{i(\mathbf{0}),j(\mathbf{R})}. \tag{27}$$

To derive the last line of Eq. (25), we have used Eq. (9), and realize that $\boldsymbol{\tau}_\mu = \boldsymbol{\tau}_i$ in the first term and $\boldsymbol{\tau}_\mu = \boldsymbol{\tau}_j$ in the second term.

Comparing Eq. (25) to (21), we arrived at the following desired result:

$$\tilde{C}^{\mu}_{i,j}(\mathbf{k}+\mathbf{q}, \mathbf{k}) = \begin{cases} \tilde{C}^{\mu(\mathbf{0})}_{i(-\mathbf{k}-\mathbf{q}),j(\mathbf{0})}, & \mu \in I \\ \tilde{C}^{\mu(\mathbf{0})}_{i(\mathbf{0}),j(\mathbf{k})}, & \mu \in J \\ 0, & \text{otherwise} \end{cases}. \tag{28}$$

Equation (28) indicates that AO triple coefficients have only two nonzero sectors, each of which depends on only one independent $\mathbf{k}$ vector instead of two. This property greatly simplifies the computation and storage of the triple coefficients, which are the key quantities in our G_0W_0 implementation.

The evaluation of two-center expansion coefficients $\tilde{C}^{\mu(\mathbf{0})}_{i(\mathbf{0}),j(\mathbf{R})}$ is described in detail in Ref. [69]. Essentially, they are determined by minimizing the self-Coulomb repulsion of the expansion error given by Eq. (24). This criterion leads to the following expression:

$$\tilde{C}^{\mu(\mathbf{0})}_{i(\mathbf{0}),j(\mathbf{R})} = \begin{cases} \sum_{\nu \in \{I, J(\mathbf{R})\}} (i(\mathbf{0}), j(\mathbf{R})|\nu)(V^{IJ})^{-1}_{\nu\mu}, & \text{for } \mu \in I \\ 0, & \text{otherwise} \end{cases}, \tag{29}$$

where $\nu \in \{I, J(\mathbf{R})\}$ means that the auxiliary function P_ν is centered either on the atom I in the original cell, or on the atom J in the cell specified by $\mathbf{R}$. Furthermore $(i, j|\nu)$ is the Coulomb repulsion between the product $\varphi_i\varphi_j$ and the ABF P_μ,

$$(i, j|\mu) = \iint d\mathbf{r}\, d\mathbf{r}' \frac{\varphi_i(\mathbf{r})\varphi_j(\mathbf{r})P_\mu(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \tag{30}$$

and V^{IJ} is a subblock of the Coulomb matrix V where the auxiliary basis indices of the entries belong to either atom I or atom J. Only an inversion of such a subblock of the Coulomb matrix is required at a time to determine the two-center expansion coefficients in Eq. (29). We note that, in Eq. (29), only two-center integrals (albeit three orbitals are involved) are required, due to the LRI approximation. Efficient algorithms exist to evaluate two-center integrals over numeric atom-centered basis functions [95–97]. All these pieces add together to make an efficient evaluation of the AO triple coefficients possible.

Furthermore, one may observe that $\tilde{C}^{\mu(\mathbf{0})}_{i(\mathbf{0}),j(\mathbf{R})} = \tilde{C}^{\mu(\mathbf{0})}_{j(\mathbf{R}),i(\mathbf{0})}$, and hence according to Eqs. (26) and (27), we have $\tilde{C}^{\mu(\mathbf{0})}_{j(\mathbf{k}),i(\mathbf{0})} =$

$\tilde{C}^{\mu(\mathbf{0})}_{i(\mathbf{0}),j(\mathbf{k})}$. It follows that Eq. (28) can be rewritten as

$$\tilde{C}^{\mu}_{i,j}(\mathbf{k}+\mathbf{q}, \mathbf{k}) = \tilde{C}^{\mu(\mathbf{0})}_{j(\mathbf{0}),i(-\mathbf{k}-\mathbf{q})} + \tilde{C}^{\mu(\mathbf{0})}_{i(\mathbf{0}),j(\mathbf{k})}. \tag{31}$$

We summarize the computational algorithm described above in terms of the flowchart in Fig. 1. The key in this algorithm is that we need only to explicitly store $\tilde{C}^{\mu(\mathbf{0})}_{i(\mathbf{0}),j(\mathbf{R})}$ and its Fourier transform $\tilde{C}^{\mu(\mathbf{0})}_{i(\mathbf{0}),j(\mathbf{k})}$. The memory-intensive $\tilde{C}^{\mu}_{i,j}(\mathbf{k}+\mathbf{q}, \mathbf{k})$ and $C^{\mu}_{m,n}(\mathbf{k}+\mathbf{q}, \mathbf{k})$ are formed on the fly when needed, within the loop over the $\mathbf{k}$ and $\mathbf{q}$ points. This efficacy of this algorithm depends on the accuracy of the LRI approximation, which further depends on the auxiliary basis set $\{P_\mu(\mathbf{r})\}$. We will discuss this issue in the next section.

The algorithm as outlined in Fig. 1 scales as $O(N^4)$ with respect to the number of one-electron basis functions and quadratically with respect to the number of $\mathbf{k}$ points. In the literature, a $O(N^3)$-scaling *GW* algorithm has been formulated based on real-space/imaginary-time representation of the response function [34,35], and is getting popular in recent years as manifested in several recent implementations within different basis set frameworks [57,98,99]. Within the NAO

---

1. Compute $\tilde{C}^{\mu(\mathbf{0})}_{i(\mathbf{0}),j(\mathbf{R})}$ using the two-center integration technique [95--97].

2. Fourier transform $\tilde{C}^{\mu(\mathbf{0})}_{i(\mathbf{0}),j(\mathbf{R})}$ to $\tilde{C}^{\mu(\mathbf{0})}_{i(\mathbf{0}),j(\mathbf{k})}$ using Eqs. (26) and (27).

3. Loop over $\mathbf{q}$

    3.1 Compute $V_{\mu\nu}(\mathbf{q})$

    3.2 Loop over $\mathbf{k}$

        a. Form $\tilde{C}^{\mu}_{i,j}(\mathbf{k}+\mathbf{q}, \mathbf{k})$ using Eq. (31).

        b. Transform $\tilde{C}^{\mu}_{i,j}(\mathbf{k}+\mathbf{q}, \mathbf{k})$ to $C^{\mu}_{m,n}(\mathbf{k}+\mathbf{q}, \mathbf{k})$ using Eq. (22).

        c. Evaluate $\chi_{0,\mu\nu}(\mathbf{q}, i\omega)$ using Eq. (11).

    End loop

    3.3 Compute $W_{0,\mu\nu}(\mathbf{q}, i\omega)$ using Eqs. (13-14).

  End loop

4. Loop over $\mathbf{k}$

    4.1 Loop over $\mathbf{q}$

        d. Rebuild $\tilde{C}^{\mu}_{i,j}(\mathbf{k}+\mathbf{q}, \mathbf{k})$ using Eq. (31), and transform to $C^{\mu}_{m,n}(\mathbf{k}+\mathbf{q}, \mathbf{k})$ using Eq. (22).

        e. Evaluate $\Sigma^{G_0W_0}_{n,\sigma}(\mathbf{k}, i\omega)$ using Eq. (17).

    End loop

  End loop

FIG. 1. Flowchart of the G_0W_0 self-energy calculation for periodic systems within the NAO framework and LRI approximation.

framework, a similar real-space algorithm can be derived, which, fostered by the LRI approximation, features an $O(N^2)$ scaling of rate-determining steps of GW self-energy calculations. The implementation work based on this algorithm is still going on. The present canonical-scaling implementation provides the reference results that one should be able to reproduce with a correct implementation of the aforementioned $O(N^2)$ real-space algorithm.

III. IMPLEMENTATION DETAILS

In the previous section, we presented the basic equations behind our implementation. Two important aspects that remain to be addressed are the procedure to construct the ABFs, which is crucial for the accuracy of the LRI approximation and the treatment of the Coulomb singularity at the Γ point. We will focus on these two aspects in this section.

A. Basis sets

In FHI-aims, the basis functions to represent the KS orbitals are given by numerically tabulated radial functions multiplied by spherical harmonics, as indicated by Eq. (19). The radial functions $u_{a,s,l}(r)$ are usually constructed to satisfy the radial Schrödinger equation (assuming the Hartree atomic unit, and abbreviating the indices $\{a, s, l\}$ by j),

$$\left[-\frac{1}{2r^2} \frac{d}{dr} r^2 \frac{d}{dr} + \frac{l(l+1)}{r^2} + v_j(r) + v_{\text{conf}}(r) \right] u_j(r)$$
$$= \epsilon_j u_j(r), \qquad (32)$$

where $v_j(r)$ is a radial potential that defines the major behavior of $u_j(r)$, whereas $v_{\text{conf}}(r)$ is a confining potential, added here in order to strictly localize $u_j(r)$ within a cutoff radius. The choice of $v_{\text{conf}}(r)$ in FHI-aims is discussed in Ref. [67]. The potential $v_j(r)$ is set to be the self-consistent free-atom potential for the atomic species s in question, to obtain the so-called *minimal* basis, consisting of core and valence wave functions of a spherically symmetric atom. Additional basis functions beyond *minimal* basis are of ionic or hydrogen type, obtained by setting $v_j(r)$ to the potential of the cations of the atomic species s, or simply the hydrogen-like potential Z/r. The effective charge Z, being fractional in general, is taken as an optimization parameter, which controls the shape and spatial extension of of the hydrogen-like orbital basis functions. These additional basis functions (mostly being hydrogen-type) are selected by optimizing the ground-state total energy of target systems and grouped into different levels. In FHI-aims, we have created two types of such hierarchical NAO basis sets. One type is the FHI-aims-2009 basis sets (often called *tier-n* basis), originally optimized for ground-state DFT calculations for symmetric dimers of varying bond lengths [67], but proved to be also useful for G_0W_0 calculations [52,59,100]. Another type is the NAO-VCC-nZ ($n = 2, 3, 4, 5$) basis sets [101], generated by optimizing the RPA total energy of free atoms. The construction of the latter type of NAO basis sets follows the "correlation consistent"

(cc) strategy of Dunning [102], and hence allows for extrapolations and suitable for correlated calculations (such as MP2, RPA, and GW). In this work, we will check the performance of both types of basis sets for periodic GW calculations.

The ABFs $\{P_\mu(\mathbf{r})\}$ are not preconstructed and optimized, but are rather formed adaptively based on a given OBS $\{\varphi_i(\mathbf{r})\}$. The detailed procedure to construct $\{P_\mu(\mathbf{r})\}$ for a given set of $\{\varphi_i(\mathbf{r})\}$ has been described in Refs. [59,69]. The essential point is that the radial functions of $\{P_\mu(\mathbf{r})\}$ are generated from the "on-site" products of the radial functions of $\{\varphi_i(\mathbf{r})\}$, and the Gram-Schmidt procedure is then used to remove the linear dependence according to a threshold θ_{orth}. The standard ABFs, generated from the OBS used in the preceding self-consistent field (SCF) calculations, are sufficiently accurate for post-DFT correlated calculations, if used in the global Coulomb-metric RI (called RI-V in Ref. [59]) framework. However, these ABFs alone are not adequate to yield the needed accuracy when used in the LRI scheme, especially for correlated methods that require unoccupied orbitals [69]. The LRI approximation employed in the present work corresponds to its nonrobust fitting formulation, and the incurred error in the two-electron Coulomb integrals are linear with respect to the expansion error of the orbital products [cf. Eq. (6)] [93,94], in contrast with the RI-V case where the error is quadratic. One way to remedy this problem is to complement the OBS with extra basis functions (called OBS+) that are used only for generating ABFs, but not in the preceding SCF calculations. It turns out that this is a very efficient way to improve upon the standard ABF set, rendering the LRI a sufficiently accurate approximation for practical calculations. In Ref. [69] it was found that adding an extra $5g$ hydrogen-like function (with $Z = 6$) to OBS results in an ABF set that is sufficiently accurate for both ground-state exact-exchange and correlated calculations (MP2 and RPA), for the molecules/clusters tested in that work. It should be noted though that the resulting ABFs for LRI can reach quite high angular momenta, e.g., up to $l = 8$ if the OBS+ basis set includes angular momenta up to $l = 4$ (see Table 12 in Ref. [69] for an example). These high angular momenta turn out to be important for the success of the LRI strategy.

In this work, we will test if the strategy of adding a $5g$ to OBS+ also works for periodic GW calculations. As will be demonstrated in Sec. IV A, for the GW case a combined $4f5g$ OBS+ can yield excellent accuracy, and will be chosen as the default setting in production calculations.

B. The Γ-point singularity treatment

For 3D systems, the $1/r$ nature of the bare Coulomb potential leads to a $1/q^2$ divergence for $q \to 0$ in reciprocal space. Within the plane wave basis, the Coulomb operator has a well-known matrix form $V_{\mathbf{G},\mathbf{G}'}(\mathbf{q}) = 4\pi \delta_{\mathbf{G},\mathbf{G}'}/|\mathbf{q} + \mathbf{G}|^2$, and the divergence is present in only the $\mathbf{G} = \mathbf{G}' = 0$ element (the so-called "head" term of a matrix with indices $\mathbf{G}$ and $\mathbf{G}'$). With the atom-centered ABFs used in this work, this divergence carries over to the matrix elements between two nodeless s-type functions ($1/q^2$ divergence), and between one nodeless s-type and one nodeless p-type functions ($1/q$

divergence). In general we can split $V_{\mu\nu}(\mathbf{q})$ into

$$V_{\mu\nu}(\mathbf{q}) = \frac{v_{\mu\nu}^{(2)}}{q^2} + \frac{v_{\mu\nu}^{(1)}}{q} + \bar{V}_{\mu\nu}(\mathbf{q}), \qquad (33)$$

where $\bar{V}_{\mu\nu}(\mathbf{q})$ is regular as $\mathbf{q} \to \mathbf{0}$, and $v_{\mu\nu}^{(2)}$ and $v_{\mu\nu}^{(1)}$ are the coefficients for elements with $1/q^2$ and $1/q$ asymptotic behaviors. Simple electrostatic analysis indicates that $v_{\mu\nu}^{(2)}$ and $v_{\mu\nu}^{(1)}$ are only nonzero for the above-noted pair of basis functions that have $1/q^2$ and $1/q$ divergences respectively.

The above diverging behavior of $V_{\mu\nu}(\mathbf{q})$ carries over to the screened Coulomb matrix $W_{0,\mu\nu}(\mathbf{q}, i\omega)$ for nonmetallic systems. Consequently, in the integration over BZ to obtain the G_0W_0 self-energy, the integrand has a $1/q^2$ divergence as $\mathbf{q}$ approaches the Γ point. This is an integrable divergence. However, when the BZ is discretized in terms of a uniform $\mathbf{k}$ mesh, the Γ point contributes to the integrated quantity, the G_0W_0 self-energy, by an amount that is proportional to the discretization length Δq, i.e., the length of the small cube enclosing the Γ point. Since $\Delta q \sim N_{\mathbf{k}}^{-1/3}$, simply neglecting the Γ point will incur a so-called "finite-size error" that decreases only linearly with $N_{\mathbf{k}}^{-1/3}$. This is a very slow convergence. Therefore, to achieve a sufficiently fast convergence of the BZ integration in Eq. (17), a special treatment of this singularity is needed. For semiconductors and insulators, the (properly treated) dielectric function is nondiverging around $\mathbf{q} = \mathbf{0}$, which means that $W_{0,\mu\nu}(\mathbf{q}, i\omega)$ has the same asymptotic behavior as $V_{\mu\nu}(\mathbf{q})$ for $\mathbf{q} \to \mathbf{0}$. Thus, the experience gained to treat the singularity of the bare Coulomb potential in periodic Hartree-Fock (HF) calculations is also useful here. In the literature, two schemes are widely adopted to deal with the Coulomb singularity in periodic HF implementations. One is the Gygi-Baldereschi scheme [103] which adds an analytically integrable compensating function to cancel the diverging term and subtracts it separately. The other is the Spencer-Alavi scheme [104] which uses a truncated Coulomb operator that is free of Coulomb singularity; yet the scheme guarantees a systematic convergence to the right limit as the number of $\mathbf{k}$ points increases. Both schemes have been implemented in our own periodic HF module [70] of the FHI-aims code. In practice, we found that a modified version of the Spencer-Alavi scheme converges faster with respect to the number of $\mathbf{k}$ points than the "compensating function" approach does. Similar observations have been made in Ref. [105], in terms of another variant of the Coulomb operator truncation scheme—the so-called Wigner-Seitz cell truncation scheme [105].

In our modified Spencer-Alavi scheme, the truncated Coulomb operator is given by

$$v^{\text{cut}}(r) = \frac{\text{erfc}(\gamma r)}{r} + \frac{1}{2}\text{erfc}\{[\ln(r) - \ln(R_{\text{cut}})]/\ln(R_w)\}$$
$$\times \frac{\text{erf}(\gamma r)}{r}, \qquad (34)$$

where the short-range part of the Coulomb potential is kept, and the long-range part is quickly suppressed beyond a cutoff radius R_{cut}. R_{cut} is chosen to be the radius of a sphere inscribed inside the BvK supercell. As the $\mathbf{k}$ point mesh gets denser, R_{cut} gradually increases and the full bare Coulomb operator is

restored. The screening parameter γ and the width parameter R_w in Eq. (34) can be tuned to achieve the best performance, but the default choice of $\gamma = 5.0/R_{\text{cut}}$ Bohr^{-1} and $R_w = 1.092$ Bohr works sufficiently well for all systems we have tested so far.

By replacing $1/|\mathbf{r} - \mathbf{r}'|$ by the truncated form $v^{\text{cut}}(|\mathbf{r} - \mathbf{r}'|)$ in Eq. (12), one obtains the truncated Coulomb matrix within the auxiliary basis, $V_{\mu\nu}^{\text{cut}}(\mathbf{q})$, which is regular for $\mathbf{q} \to \mathbf{0}$. A corresponding truncated screened Coulomb matrix $W_0^{\text{cut}}(\mathbf{q}, i\omega)$ can then be defined via Eq. (14) by replacing the full $V(\mathbf{q})$ matrix by the truncated one $V^{\text{cut}}(\mathbf{q})$ in the numerator. As described below, in doing so, one should be careful that the symmetrized dielectric function in the denominator of Eq. (14) should not be affected by the truncation procedure, i.e., it should still be determined using the full Coulomb operator. For semiconductors and insulators, which are our concerns in the present work, the symmetrized dielectric function $\tilde{\varepsilon}$ matrix is finite and invertible everywhere in the BZ. The $\tilde{\varepsilon}$ matrix can be directly computed using Eq. (13) for all $\mathbf{q}$ points except at $\mathbf{q} = 0$. Because of the diverging behavior of $V_{\mu\nu}(\mathbf{q})$ for $\mathbf{q} \to 0$, as indicated by Eq. (33), the asymptotic behavior of the χ_0 matrix [Eq. (11)] needs to be taken care of in order to cancel the divergence in the V matrix, similar to what is done in the plane-wave representation [91,92].

To this end, we adopt the scheme widely used in the context of the LAPW framework, which represents the dielectric function in terms of the eigenvectors of the Coulomb matrix [38–40]. At $\mathbf{q} = 0$, instead of diagonalizing the diverging full Coulomb matrix $V(\mathbf{k} = 0)$, we diagonalize the truncated Coulomb matrix $V^{\text{cut}}(\mathbf{k} = 0)$,

$$\sum_{\nu} V_{\mu,\nu}^{\text{cut}}(\mathbf{k} = 0)X_{\nu,\lambda} = X_{\mu,\lambda}v_{\lambda}, \qquad (35)$$

where v_{λ} and $X_{\mu,\lambda}$ are its eigenvalues and its eigenvectors, which are real-valued for $\mathbf{k} = 0$. Within its broad eigenspectrum, there is one eigenstate standing out, with an eigenvalue that is significantly larger than all others, corresponding to the $G = 0$ plane wave (i.e., the constant $1/\sqrt{\Omega}$ with Ω being the volume of the crystal). We shall denote this eigenstate as $\lambda = 1$, and order the rest eigenstates according to their eigenvalues (v_{λ} with $\lambda > 1$) in an energetically descending manner. Naturally, $v_{\lambda=1}$ goes to infinity as $N_{\mathbf{k}}$ increases, whereby V^{cut} approaches the full V. Analyzing the components of the eigenvector $X_{\mu,1}$ reveals that this state has a predominant contribution from the nodeless s functions— the only ABFs representing nonzero net charges—consistent with the nature of the $\mathbf{G} = 0$ plane wave. For the Coulomb matrix with a small q, we can diagonalize both the full Coulomb matrix $V(\mathbf{q})$ and its regular part $\bar{V}(\mathbf{q})$ [cf. Eq. (33)], and find that their eigenvectors are nearly the same. This is because the diverging part of the Coulomb matrix originates entirely from the $\lambda = 1$ eigenvector, and removing this part from the Coulomb matrix amounts to shifting the value of v_1 downwards, but not changing the eigenvector $X_{\mu,1}$. Based on the above understanding, and inspired by the prescription of the LAPW basis set [38–40,106], we arrive at the following expression for the symmetrized dielectric function matrix within the Coulomb eigenvector basis

representation:

$$\tilde{\varepsilon}_{\lambda,\lambda'}(\mathbf{q} \to \mathbf{0}, i\omega)$$

$$= \begin{cases} 1 - \dfrac{4\pi}{\Omega} \sum_{\mathbf{k},\sigma} \left\{ \sum_n \dfrac{f'(\epsilon_{n,\sigma}^{\mathbf{k}}) |\mathbf{p}_{n,n,\sigma}^{\mathbf{k}} \cdot \hat{\mathbf{q}}|^2}{\omega^2} + \sum_{m<n} \dfrac{2(f_{m,\sigma}^{\mathbf{k}} - f_{n,\sigma}^{\mathbf{k}}) |\mathbf{p}_{m,n,\sigma}^{\mathbf{k}} \cdot \hat{\mathbf{q}}|^2}{\left[(\epsilon_{m,\sigma}^{\mathbf{k}} - \epsilon_{n,\sigma}^{\mathbf{k}})^2 + \omega^2\right](\epsilon_{m,\sigma}^{\mathbf{k}} - \epsilon_{n,\sigma}^{\mathbf{k}})} \right\}, & \lambda = \lambda' = 1 \\[12pt] -\sqrt{\dfrac{4\pi}{\Omega N_{\mathbf{k}}}} \sum_{\mathbf{k},\sigma} \sum_{m<n} \sum_\nu \dfrac{f_{m,\sigma}^{\mathbf{k}} - f_{n,\sigma}^{\mathbf{k}}}{\epsilon_{m,\sigma}^{\mathbf{k}} - \epsilon_{n,\sigma}^{\mathbf{k}}} \left[\dfrac{(\mathbf{p}_{m,n,\sigma}^{\mathbf{k}} \cdot \hat{\mathbf{q}}) C_{n,m,\sigma}^\nu(\mathbf{k},\mathbf{k})}{\epsilon_{m,\sigma}^{\mathbf{k}} - \epsilon_{n,\sigma}^{\mathbf{k}} - i\omega} + \text{c.c.} \right] X_{\nu,\lambda'} \sqrt{v_{\lambda'}}, & \lambda = 1, \lambda' > 1 \\[12pt] -\sqrt{\dfrac{4\pi}{\Omega N_{\mathbf{k}}}} \sum_{\mathbf{k},\sigma} \sum_{m<n} \sum_\mu \sqrt{v_\lambda} X_{\mu,\lambda} \dfrac{f_{m,\sigma}^{\mathbf{k}} - f_{n,\sigma}^{\mathbf{k}}}{\epsilon_{m,\sigma}^{\mathbf{k}} - \epsilon_{n,\sigma}^{\mathbf{k}}} \left[\dfrac{C_{m,n,\sigma}^\mu(\mathbf{k},\mathbf{k}) (\mathbf{p}_{n,m,\sigma}^{\mathbf{k}} \cdot \hat{\mathbf{q}})}{\epsilon_{m,\sigma}^{\mathbf{k}} - \epsilon_{n,\sigma}^{\mathbf{k}} - i\omega} + \text{c.c.} \right], & \lambda > 1, \lambda' = 1 \\[12pt] \delta_{\lambda,\lambda'} - \dfrac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k},\sigma} \sum_{m<n} \sum_{\mu,\nu} \sqrt{v_\lambda} X_{\mu,\lambda} (f_{m,\sigma}^{\mathbf{k}} - f_{n,\sigma}^{\mathbf{k}}) \left[\dfrac{C_{m,n,\sigma}^\mu(\mathbf{k},\mathbf{k}) C_{n,m,\sigma}^\nu(\mathbf{k},\mathbf{k})}{\epsilon_{m,\sigma}^{\mathbf{k}} - \epsilon_{n,\sigma}^{\mathbf{k}} - \omega} + \text{c.c.} \right] X_{\nu,\lambda'} \sqrt{v_{\lambda'}}, & \lambda > 1, \lambda' > 1 \end{cases}, \quad (36)$$

where $\mathbf{p}_{m,n,\sigma}^{\mathbf{k}} = \langle \psi_{m,\sigma}^{\mathbf{k}} | \hat{p} | \psi_{n,\sigma}^{\mathbf{k}} \rangle = -i \langle \psi_{m,\sigma}^{\mathbf{k}} | \boldsymbol{\nabla} | \psi_{n,\sigma}^{\mathbf{k}} \rangle$ is the so-called momentum matrix, $\hat{\mathbf{q}} = \mathbf{q}/q$ is the unit vector along the direction of $\mathbf{q}$, and c.c. denotes the complex conjugate. Furthermore, $f'(\epsilon)$ is the energy derivative of the Fermi-Dirac function,

$$f'(\epsilon) = \delta f(\epsilon)/\delta\epsilon = -\dfrac{\exp[(\epsilon - \mu)/\Delta]}{\Delta\{1 + \exp[(\epsilon - \mu)/\Delta]\}^2}, \quad (37)$$

which becomes a δ-function, i.e., $-\delta(\epsilon - \mu)$, when the broadening parameter Δ (introduced to stabilize the calculations for metals or narrow-gap insulators) approaches zero. Equation (36) indicates that the head and wing terms [$\lambda = 1$ or $\lambda' = 1$ in Eq. (36)] of the dielectric function matrix in general depend on the direction along which $\mathbf{q}$ approaches zero. This is indeed the case for anisotropic systems. The first term in the first line of Eq. (36) corresponds to the so-called intraband contribution, which is present only for metals. For insulators and semiconductors, this term is zero and doesn't need to be considered.

Now we have a computable formalism for the dielectric function in the basis representation of Coulomb eigenvectors. After its computation, we can transform it back to the basis of ABFs,

$$\tilde{\varepsilon}_{\mu,\nu}(\mathbf{q} \to \mathbf{0}, i\omega) = \sum_{\lambda,\lambda'} X_{\mu,\lambda} . \tilde{\varepsilon}_{\lambda,\lambda'}(\mathbf{q} \to \mathbf{0}, i\omega) X_{\nu,\lambda'}. \quad (38)$$

The matrix form of the truncated screened Coulomb interaction $W_{0,\mu,\nu}^{\text{cut}}(q, i\omega)$ can then be computed in the entire BZ, and the GW self-energy can be calculated by standard BZ sampling techniques.

Note that our scheme to deal with the Coulomb singularity in G_0W_0 calculations as outlined above is different from what is usually done in the literature. In the usual practice [33,40,43,107,108], once the symmetrized dielectric function is properly treated at $\mathbf{q} = 0$, one can compute $\tilde{\varepsilon}^{-1}$ by performing a blockwise inversion and proceed to obtain the screened Coulomb interaction $W_0(\mathbf{q}, i\omega)$. $W_0(\mathbf{q}, i\omega)$ has a similar $1/q^2$ singularity behavior as the bare Coulomb interaction $V(\mathbf{q})$ for the so-called "head" term, and additionally a $1/q$ singular behavior for the wing terms. One can then subtract two analytic compensating functions from the integrand to smooth out these singular behaviors a $\mathbf{q} \to 0$, and the BZ integration over these compensating functions can be done separately in an analytical way. In the end, the entire procedure boils down to

performing an usual BZ summation (whereby the Γ-point can be omitted), and then adding two correction terms afterwards. Thus, the usual procedure can be viewed as the G_0W_0 analogy of the Gygi-Balderenschi scheme in HF calculations, whereas our above-described scheme follows the spirit of the Alavi-Spencer scheme. A systematic comparison of the performance of these two schemes is of high academic interest, but goes beyond the scope of the present work.

IV. CONVERGENCE TESTS

Let us now investigate the precision and convergence behavior of our G_0W_0 implementation with respect to the numerical settings, including the ABF basis set, the one-electron basis set, and the $\mathbf{k}$ point summation in 1BZ.

A. Convergence test with respect to the number and shapes of ABFs

As discussed in Sec. II C, our G_0W_0 implementation relies on the LRI approximation, whose accuracy further depends on the quality of the ABF basis set. In FHI-aims, the standard ABFs are constructed on the fly from the one-electron OBS employed in the preceding (g)KS calculations. As demonstrated in Ref. [69], within LRI, typically a larger number of ABFs, especially those with higher angular momenta, is needed to achieve similar accuracy as in the standard RI-V scheme [59]. As mentioned already in Sec. III A, one practical way to improve the accuracy of LRI is to complement the OBS with additional functions of high angular momenta. The crucial point is that these additional functions are used only to construct ABFs, but not used in Eq. (7) or (18) to expand the eigen-orbitals. Namely, they don't enter the preceding self-consistent KS-DFT calculations. In this way, the orbital products $\psi_{m,\sigma}^*(\mathbf{r})\psi_{n,\sigma}(\mathbf{r})$ on the left side of Eq. (6) does not change, but the $\{P_\mu\}$ set on the right side gets increased. This can lead to improved accuracy of the expansion in Eq. (6), and thus the final results of the ground-state energy and/or quasiparticle band structure. In the FHI-aims input file, these additional functions are labeled with a tag *for_aux*, signifying that these functions are used only to generate auxiliary functions. We follow the nomenclature of Ref. [69] by terming these additional *for_aux* functions as enhanced orbital basis set (OBS+).

In Ref. [69], it has been shown that, for an OBS that contains at least up to f functions, adding an additional $5g$ hydrogenic function to the OBS+ can yield an ABF set that is sufficiently accurate for both HF and correlated methods like MP2 and RPA in molecular calculations. The shape and spatial extension of the $5g$ hydrogenic function $u(r)$ depend on the effective nuclear potential, $v(r) = Z/r$, with Z being an effective charge, which governs the spatial extent of the solution of the radial Schrödinger equation [Eq. (32)]. The smaller the value Z is, the more extended the radial function $u(r)$. It was found [69] that the HF or MP2 results are not very sensitive to the precise value of Z, and an additional $5g$ function defined by $Z = 6.0$ is usually sufficiently good, yielding an accuracy that is comparable to the standard RI-V approximation.

Here we will examine the influence of the additional functions included in OBS+ on the band gaps of the periodic G_0W_0 calculations. To this end, we take Si and MgO crystals as the benchmark systems. The experimental lattice parameters ($a = 5.431$ Å for Si and 4.213 Å for MgO) are used in all calculations. Si is a covalently bonded semiconductor with a medium, indirect band gap, whereas MgO is an ionic crystal with a wide, direct band gap. Thus the two systems can be taken as representative examples for semiconductors and insulators, well suited for benchmark purposes.

In Fig. 2 we present the calculated G_0W_0 band gaps as a function of effective charge Z, for different *for_aux* functions and their combinations included in OBS+. For the tests to be systematic, we have here checked not only the influence of $5g$ functions but also of $4f$ and $6h$ functions. The reference state for G_0W_0 calculations in Fig. 2 is generated from KS-DFT under the generalized gradient approximation (GGA) of Perdew, Burke, and Ernzerhof (PBE). The FHI-aims *tier* 2 basis and a $8 \times 8 \times 8$ **k** grid were used in the PBE calculation. Note that the *tier* 2 basis set contains basis functions up to g for Si and O, and up to f for Mg. The G_0W_0 band gap was determined from full band structure calculations along high symmetry lines in the BZ.

The *for_aux* functions are added in the following manner. We first add a hydrogen-like $4f$ function, generated with an effective potential Z/r; by varying the value of Z, we check the influence of the shape of the added $4f$ function on the calculated G_0W_0 band gap. Next we fix the $4f$ function with an "optimal" Z value, and add an additional $5g$ function governed by its own Z value. Then we repeat this process by fixing both the $4f$ and $5g$ functions, and check the influence of an additional $6h$ function of varying spatial extent. Our rule to choose the "optimal" Z value at each step is somewhat arbitrary, but here we pick the value in a window of $1.0 <= Z <= 8.0$ that gives the biggest increase of the band gap. Figure 2(a) shows that, for Si, adding a $4f$ *for_aux* function can enlarge the calculated G_0W_0 gap roughly from 0.01 eV to 0.03 eV, depending on the chosen Z value. Smaller Z values (more extended $4f$ functions) tend to bring bigger corrections. Now, fixing the $4f$ function at $Z = 1.0$, and adding a $5g$ function in addition with varying Z [denoted as $4f(1.0) + 5g$], one can see that the band gap increases by about 2 meV, regardless the value of Z. This is because once the $4f$ function is added, the largest part of the LRI error is removed, and the remaining error is not sensitive to the shape of the ABFs anymore. This is
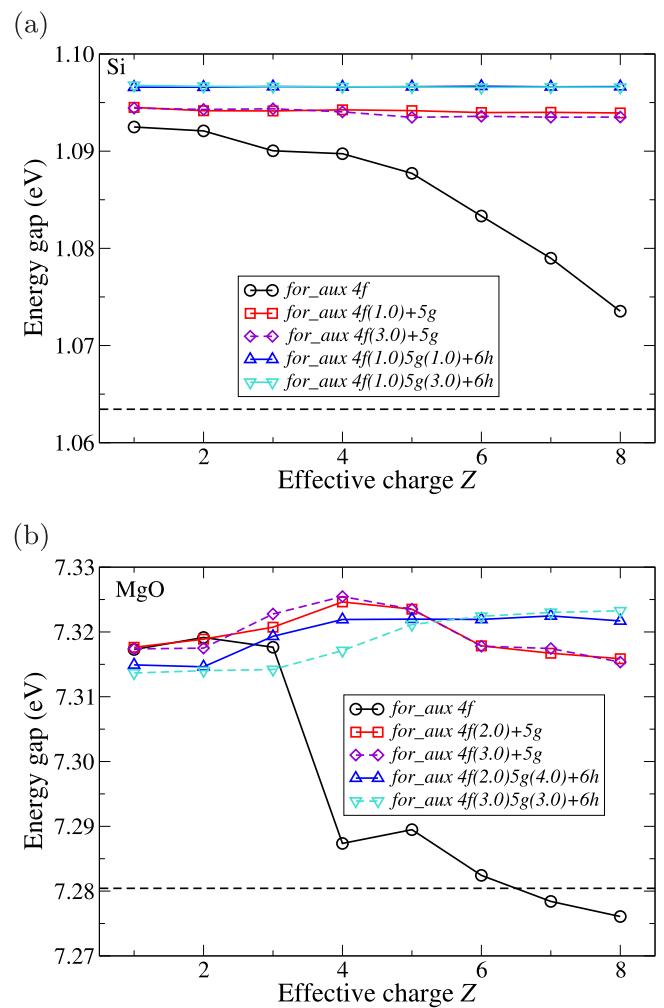

FIG. 2. The G_0W_0@PBE band gaps for Si (a) and MgO (b) as a function of the effective charge Z of the added hydrogen-like *for_aux* function used to generate additional ABFs. The legend $4f(1.0) + 5g$ means that in the OBS+, a $4f$ function with effective charge $Z = 1.0$ and a $5g$ function with varying effective charge Z are added. Likewise, $4f(2.0)5g(4.0) + 6h$ means a $4f$ with $Z = 2.0$, a $5g$ function with $Z = 4.0$, and a $6h$ function with varying Z. The dashed lines illustrate the behavior if the default Z parameters for $4f/5g$ *for_aux* functions are used. The FHI-aims *tier* 2 basis and a $8 \times 8 \times 8$ **k** grid were used in the preceding PBE calculations. The G_0W_0 band gaps are determined from full band structure calculations. The horizontal dashed lines mark the results without including any *for_aux* functions.

exactly the kind of effect we would like to achieve. Similarly, fixing $4f$ and $5g$ functions both at $Z = 1$, and adding one $6h$ function [denoted as $4f(1.0)5g(1.0) + 6h$], one gets a further increase of 2 meV of the band gap, independent of the Z value of the $6h$ function. Such a convergence behavior with respect to the *for_aux* functions in OBS+ strongly suggests the remaining error arising from LRI is minor, and the G_0W_0 band gap for Si is well converged within 0.01 eV with respect to the ABFs.

Similarly, for MgO, the change of the G_0W_0 band gap after adding a single $4f$ *for_aux* function (to both elements) is quite sensitive to the Z value. As can be seen in Fig. 2(b), the $4f$

functions with smaller Z values lead to bigger increases of the band gap (as large as 0.04 eV), while those with larger Z values bring smaller increases (or even slight decreases) of the band gap. Fixing the $4f$ function at $Z = 2.0$ and adding a $5g$ function with varying Z, the band gap shows very little change for small Z's and a slightly larger increase for big Z values. The overall effect is that the dependence of the obtained G_0W_0 band gap on the shape of the $5g$ function is much reduced, although a remaining variation within a window of 0.01 eV can still be observed. Finally, fixing $4f$ and $5g$ functions, and adding a $6h$ function with varying Z, the dependence of the obtained band gap on the Z value of $6h$ function is further reduced. In contrast to the case of Si, for MgO the calculated band gap does not always get increased upon adding more *for_aux* functions. Instead, a slight decrease of the band gap can happen occasionally, indicating a more complex convergence behavior of the G_0W_0 calculation with respect to the ABFs. Nevertheless, the overall variation of the band gap values upon including additional $5g$ or $6h$ *for_aux* functions with varying Z never exceed a range of 0.02 eV. Such an uncertainty does not affect our subsequent convergence tests with respect to other computational parameters and the final benchmark calculations.

In the above tests, the chosen "optimal" Z values for $4f/5g$ *for_aux* functions are different for Si and MgO, but importantly, the above discussed convergence behavior, as well as the final results are not sensitive to the actual Z values over a wide range. For example, using $4f(3.0)5g(3.0)$ instead of the "optimal" $4f(1.0)5g(1.0)$ for Si and $4f(2.0)5g(4.0)$ for MgO, one finds a difference of only 0.2 meV for Si and 2 meV for MgO in the calculated G_0W_0@PBE band gap. The dashed lines in Fig. 2 illustrate what happens if one starts with $4f(3.0)$ and $4f(3.0)5g(3.0)$ *for_aux* functions. Furthermore, adding $6h$ *for_aux* functions to the OBS+ leads to a factor of 2 increase of the computational cost but the accuracy gain is minor (the band gap changes below 0.01 eV). It should be noted that, in our present implementation, the computational cost is governed by only the size of the auxiliary basis set, but not the shape of the ABFs generated from the *for_aux* functions. That is, the computational cost is not affected by the Z values of the *for_aux* functions. Under such circumstances, instead of using different settings for different systems, we use a universal $4f(3.0)5g(3.0)$ *for_aux* basis setting in the following convergence tests and benchmark calculations for all materials.

Finally, to demonstrate that the error incurred by LRI has indeed been made insignificantly small by adding *for_aux* functions as outlined above, we compare our calculated full G_0W_0 band structure to independent reference results as obtained by the all-electron, LAPW-based FHI-gap code [40,109], using MgO as an example. In Fig. 3 the calculated band structure of MgO are presented for both PBE and G_0W_0@PBE. The calculations employ the FHI-aims *tier* 2 one-electron basis set, a fairly dense $8 \times 8 \times 8$ **k** grid, and the above-noted $4f(3.0)5g(3.0)$ *for_aux* functions. For comparison, the PBE and G_0W_0@PBE band structures, as obtained respectively by the LAPW-based WIEN2k code [110] and the FHI-gap code, are shown as blue dashed lines in Fig. 3. With its recent extension to complement the standard LAPW basis set with high-energy local orbitals (HLOs) [111] in its GW calculations, the FHI-gap code, interfaced with WIEN2k
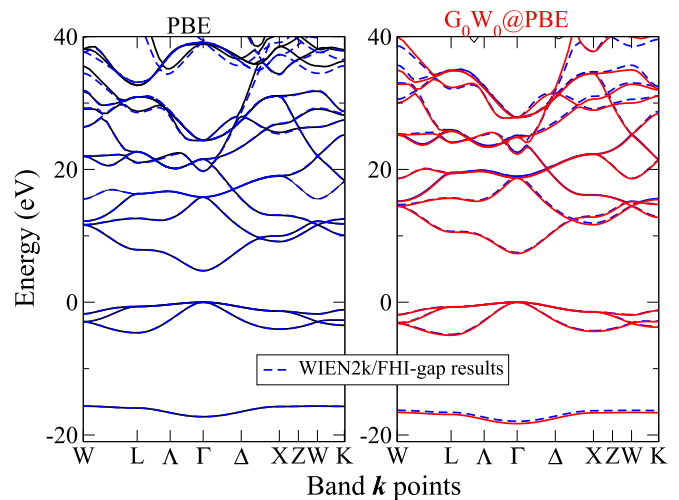

FIG. 3. The PBE (left panel) and G_0W_0@PBE (right panel) band structures for MgO. Full lines are FHI-aims results (this work), obtained using *tier* 2 basis set, the $4f(3.0)5g(3.0)$ *for_aux* functions, and a $8 \times 8 \times 8$ **k** point mesh. The blue dashed lines in the right panel are the G_0W_0 band structure produced by the LAPW-based FHI-gap [40,109] code, based on the PBE band structure determined by WIEN2k (blue dashed lines in the left panel).

[110], has been shown to deliver highly accurate GW band gaps [90] for a range of semiconductors and insulators. These include a set of 24 semiconductors and insulators that are typically used to benchmark the accuracy of theoretical approaches to electronic band structure of materials [90], as well as copper and silver halides (CuX and AgX with X=Cl, Br, I) [112], and several d- and f-electron oxides [113]. From Fig. 3, it can be seen that the FHI-aims results agree with those of FHI-gap very well over the entire BZ, not only for the valence-band maximum (VBM) and conduction-band minimum (CBM), but also for quasiparticle energy levels much higher and lower in energy. Such close agreement between two very different G_0W_0 computational frameworks is remarkable. In fact, excellent agreement has also been achieved at the PBE level between the two all-electron codes – FHI-aims and WIEN2k, as can be seen in the left panel of Fig. 3. More details can be found in Refs. [72,74]. Furthermore, from Fig. 3, we see that with the standard NAO *tier* 2 basis set, we can describe not only the occupied quasiparticle bands, but also the unoccupied bands up to 40 eV with acceptable accuracy.

As shown above, compared to Si, MgO is a somewhat more challenging system to control the error associated with LRI approximation, used in our G_0W_0 implementation. Yet, an excellent agreement between our NAO-based implementation [with an universal $4f(3.0)5g(3.0)$ *for_aux* function setting] and the LAPW-based implementation – here the FHI-gap code can be achieved. A systematic benchmark study for a range of materials will be given in Sec. V. In the remaining part of this section, we continue to examine how our our G_0W_0 implementation converges with other numerical settings.

B. Convergence test with respect to k points

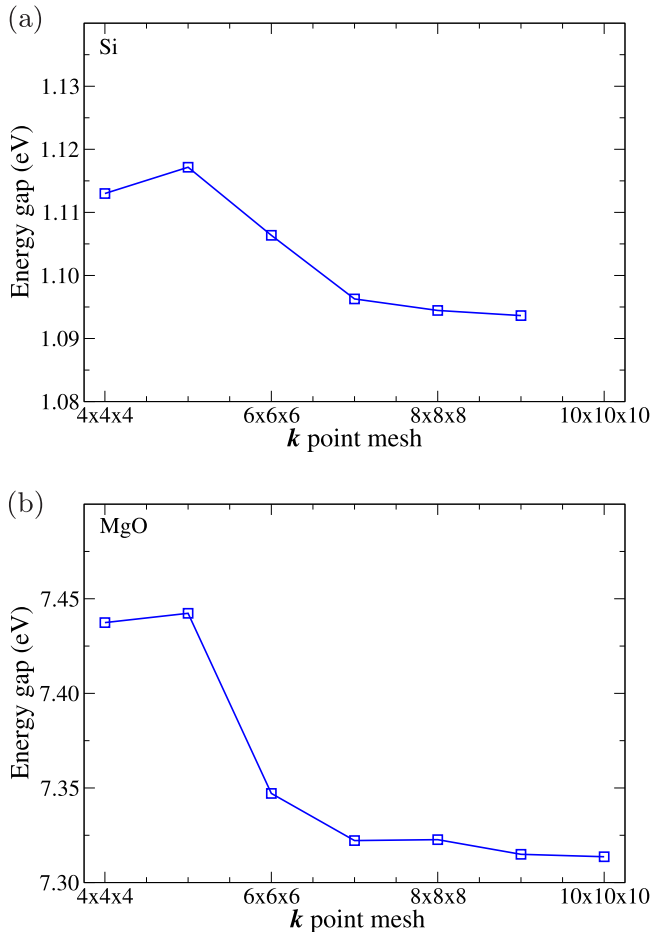In the above subsection, we discussed the convergence behavior of our G_0W_0 results with respect to the ABFs (more

FIG. 4. The G_0W_0@PBE band gaps for Si (a) and MgO (b) as a function of the **k** point mesh in the BZ summation. A uniform mesh including the Γ point is used. The FHI-aims *tier* 2 is employed for the one-electron basis set, and the *for_aux* $4f(3.0)5g(3.0)$ functions are used to generate additional ABFs.

precisely the *for_aux* functions used to generate additional ABFs), for a given **k** point mesh. With errors arising from the LRI approximation under control, we next examine how our G_0W_0 calculations converge with respect to the **k** point summation in evaluating the self-energy.

As discussed in Sec. III B, the summand in Eq. (17) has an integrable divergence around $\mathbf{q} = 0$, and the convergence behavior of a finite summation over **q** crucially depends on how we treat this singularity. In Fig. 4 we present the G_0W_0@PBE band gap for Si and MgO as a function of the **k** point mesh in the BZ summation. Calculations are done with *tier* 2 one-electron basis set, and $4f(3.0)5g(3.0)$ *for_aux* functions. For Si, the calculated band gap varies smoothly as the number of **k** points increases. Initially, it drops for increasing the **k** point density, but quickly saturates at around 1.09 eV for **k** meshes of $7 \times 7 \times 7$ and denser. For MgO, the convergence of the band gap with respect to the **k** point mesh follows a similar behavior. The band gap essentially remains constant for a **k** point mesh beyond $7 \times 7 \times 7$, and saturates at a value between 7.31 and 7.32 eV, with a tiny remaining variation of 0.01 eV. The origin of this additional complication is likely related to our special way of treating

the Γ point singularity: the truncated bare Coulomb operator together with an explicit evaluation of the dielectric function at $\mathbf{q} = 0$ within the Coulomb eigenvector basis representation, as described in Sec. III B. Investigating the origin of this issue is beyond the scope of this paper, which focuses on the numerically converged **k** grid setting.

Below we will use $8 \times 8 \times 8$ **k** point grid for further convergence studies with respect to other parameters and final benchmark calculations for crystals with a zinc-blende structure. For wurzite structure we use a $8 \times 8 \times 5$ **k** grid instead, due to the larger size of the unit cell along the z direction.

C. Convergence test with respect to one-electron orbital basis sets

In the above two subsections, we have examined the convergence behavior of our G_0W_0 implementation with respect to the *for_aux* functions and **k** point meshes for Si and MgO. We demonstrate that the numerical errors stemming from the LRI approximation and finite **k** point sampling can be well controlled and the uncertainties of the calculated G_0W_0 band gap due to these factors are within 0.01–0.02 eV. In this section we will check how our G_0W_0 results converge with respect to the one-electron OBS. Here we emphasize again that the one-electron OBS discussed in this subsection should be distinguished from the *for_aux* functions (included in OBS+) discussed in Sec. IV A. While the former is used to expand the SCF molecular eigen-orbitals that enter the subsequent G_0W_0 calculations, the latter (OBS+) is employed only to generate additional ABFs so as to reduce the numerical error incurred by the LRI approximation, instead of improving the description of one-electron eigen-orbitals.

The convergence to the complete basis set (CBS) limit using local atomic orbitals for correlated methods has long been considered a challenging problem. In quantum chemistry, experience has been gained to reach the CBS limit when using *correlation consistent* or *balanced* GTO basis sets [102,114] combined with the Helgaker extrapolation procedure [115]. Such a procedure works nicely for small molecules, but cannot be directly applied to solids, due to a tendency to encounter linear dependence when using the standard GTOs in close packed structures. In this regard, NAOs are expected to be better behaved due to their strict locality in real space. Here we will first examine how the G_0W_0 band gaps converge with respect to the standard NAOs utilized in FHI-aims, and then further investigate the influence of highly localized Slater-type orbitals (STOs).

1. G_0W_0 calculations with one-electron NAO basis sets

For NAO basis sets, remarkable all-electron precision can be obtained for ground-state DFT calculations involving only occupied states [67,73]. However, similar to the GTO case, the situation gets more involved for correlated calculations, like MP2 and RPA. The FHI-aims-2009 ("*tier*") basis sets, originally developed for ground-state DFT calculations, can yield acceptably accurate results for MP2 and RPA binding energies if the basis set superposition errors (BSSEs) are corrected [59]. Better accuracy can be achieved when using NAO-VCC-nZ basis sets [101], via which results can be extrapolated to the CBS limit using a two-point extrapolation procedure

[115]. However, it was found that the original NAO-VCC-nZ basis sets developed for molecules are not optimal for close-packed solids, due to their relatively larger radial extent (compared to the FHI-aims-2009 basis sets). Zhang *et al.* [116] reoptimized these basis sets by removing the so-called "enhanced minimal basis" and tightening the cutoff radius of the basis functions. The resultant basis sets, called localized NAO-VCC-nZ (loc-NAO-VCC-nZ) here, have been shown in Ref. [116] to be able to yield accurate MP2 and RPA binding energies for simple solids, with an appropriate extrapolation procedure.

The basis-set convergence of G_0W_0 energy levels is again different from MP2 and RPA for NAO basis sets. The FHI-aims-2009 basis sets are usually used in FHI-aims G_0W_0 calculations for molecules. When using the *tier* 4 basis set plus some additional diffuse GTOs, the G_0W_0 calculations yield results comparable to those obtained with a very large aug-cc-pV5Z GTO basis sets [59,63,100]. In the G_0W_0 case, the advantage of NAO-VCC-nZ basis sets is therefore less obvious, with a typically comparable or even slower convergence behavior. Also, a reliable and easy-to-use procedure to extrapolate G_0W_0 results to the CBS limit is not yet established. The different convergence behavior of NAOs for G_0W_0 and for MP2/RPA is due to the fact that these methods are used to calculate different energetic properties. While G_0W_0 deals with individual states for which the error due to inaccurate unoccupied orbitals is not significant, MP2 and RPA calculate total energies for which relatively small errors for individual states can accumulate to an unmanageable level.

Now we are at a stage to test the convergence behavior of both loc-NAO-VCC-nZ and FHI-aims-2009 basis sets for periodic G_0W_0 calculations. In Fig. 5 we present the calculated G_0W_0 band gaps for Si and MgO as a function of the basis set size. The **k** point mesh is fixed at $8 \times 8 \times 8$ and the $4f(3.0)5g(3.0)$ *for_aux* functions are used to generate additional ABFs. For Si, the band gap only slightly decreases from 1.098 eV to 1.091 eV from *tier* 1 to *tier* 3. On the other hand, when the loc-NAO-VCC-nZ basis sets are used, the band gap starts with a high value of 1.178 eV with the loc-NAO-VCC-2Z basis, but quickly drops to 1.083 eV for the loc-NAO-VCC-3Z basis set and slightly increases back to 1.093 eV for the loc-NAO-VCC-4Z basis set. Thus, with the largest NAO-VCC-4Z and *tier* 3 basis sets, we can achieve a satisfying agreement within 2 meV for the G_0W_0@PBE band gap. It is also interesting to see that the *tier* 1 and *tier* 2 basis sets can already yield band gaps that differ from the *tier* 3 result only by a few meV. From Fig. 5(a) it can be seen that only the loc-NAO-VCC-2Z basis set appears to be out of the general trend. This is because, by eliminating the "enhanced minimal basis," the loc-NAO-VCC-2Z is not yet sufficient to accurately describe the occupied states in the preceding SCF calculations. Nevertheless, the overall convergence behavior indicates that, using either the FHI-aims-2009 or loc-NAO-VCC-nZ basis sets, we can safely converge the G_0W_0 gap for Si within 0.01 eV.

For MgO, using the FHI-aims-2009 basis sets, the G_0W_0@PBE gap increases from 7.31 eV to 7.35 eV. When the loc-NAO-VCC-nZ basis sets are used, the calculated band gap quickly drops by about 0.4 eV from loc-NAO-VCC-2Z to loc-NAO-VCC-3Z. Further increasing the basis set to
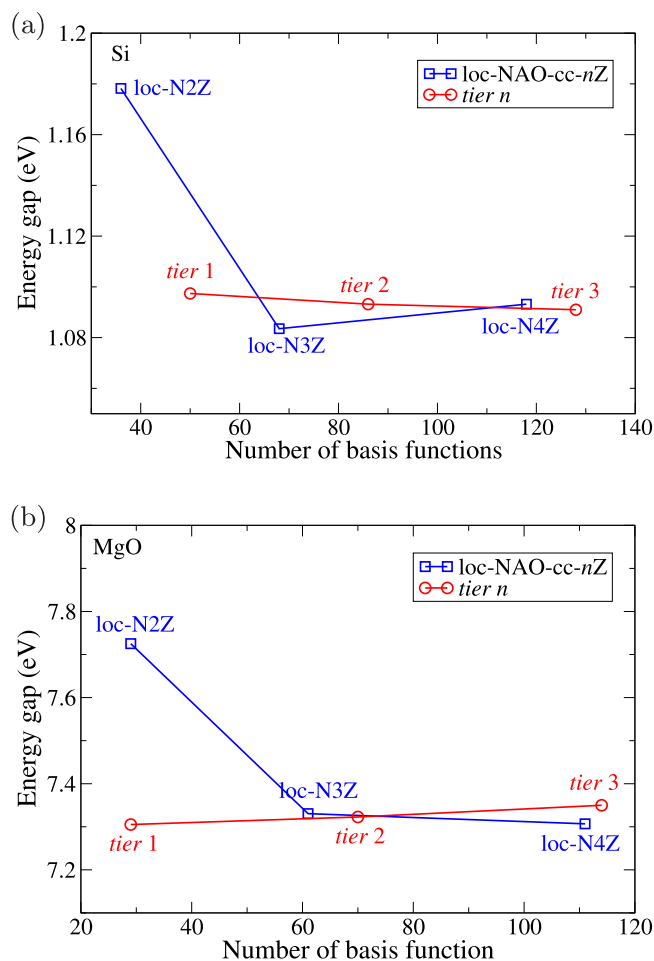

FIG. 5. The G_0W_0@PBE band gaps for Si (a) and MgO (b) as a function of the size of the one-electron basis set per formula unit cell. Two types of hierarchical basis sets, the loc-NAO-VCC-nZ (blue squares) and the FHI-aims-2009 ("*tier*") basis sets (red circles), are used for the convergence test calculations. The "loc-NAO-VCC-nZ" is abbreviated as "loc-NnZ" to label the data points. A combined set of $4f(3.0)5g(3.0)$ hydrogen-like *for_aux* functions and a $8 \times 8 \times 8$ **k** grid were used in the calculations.

loc-NAO-VCC-4Z results in a slight decrease of the band gap from 7.33 to 7.31 eV. As such, the band gap values obtained with the *tier* 2 and loc-NAO-VCC-nZ basis sets are already very close, but the difference becomes slightly larger when both types of basis sets get further increased. This behavior may reflect the presently achievable level of basis set convergence, since the basis set sizes of *tier* 3 and loc-NAO-VCC-4Z are already rather large and the basis functions in a close-packed solids becomes linear dependent. In practical calculations, if the linear dependence becomes severe, we perform singular-value decomposition (SVD) to remove the redundant components, specifically by eliminating the eigenfunctions of the overlap matrix below a threshold of 10^{-4}. We note that the SVD procedure, while making it possible to run calculations with large and linearly dependent basis sets, introduces some numerical uncertainty since the final G_0W_0 result will noticeably depend on the choice of the threshold value. This issue is more pronounced for MgO than

Si, since MgO has a smaller lattice constant and Mg has a larger basis cutoff radius. For example, when *tier* 3 basis set is used, the G_0W_0@PBE gap for MgO may range from 7.22 eV to 7.44 eV, depending on the choice of the SVD threshold. Such an uncertainty is too large to be acceptable. Thus for MgO, we take our *tier* 2 result (7.32 eV) as the most accurate G_0W_0@PBE gap value that our present numerical framework could offer.

2. Complementing the one-electron NAOs with highly localized STOs

With the standard NAOs, we face the problem that the quality of the convergence of the G_0W_0 results with respect to one-electron OBS cannot be rigorously assessed, because systematically increasing the size of NAOs to the complete basis set limit is difficult due to their nonorthogonal nature. Already for MgO, the reliability of the results obtained with *tier* 3 or loc-NAO-VCC-4Z basis sets is limited by the numerical instability arising from the linear dependence of the basis functions. In the LAPW framework, it has been observed that adding high-energy localized orbitals within the muffin-tin sphere to the standard LAPW basis can significantly improve the band gap values for materials such as ZnO [41,90,117]. Inspired by this experience, we test here the possible impact by complementing the standard NAOs with highly localized orbitals. As a first numerical test, highly localized Slater-type orbitals (STOs) are used. The reason to use STOs instead of NAOs here is that it is easier to control the spatial extent of STOs by a single parameter. Furthermore, they can be deployed in an even-tempered way to systematically span the Hilbert space. The even-tempered STOs [118] are given by

$$\phi_{\kappa,n,l,m}^{\text{STO}}(\mathbf{r}) = \mathcal{N} r^{n-1} e^{-\zeta_{l,\kappa} r} Y_{lm}(\hat{\mathbf{r}}), \tag{39}$$

where $\mathcal{N} = (2\zeta)^n \sqrt{2\zeta/2n!}$ is a normalization factor, and

$$\zeta_{l,\kappa+1} = \alpha_l \beta_l^{\kappa-1} 2\zeta_{l,\kappa}, \kappa = 1, 2, 3, \ldots. \tag{40}$$

That is, the exponents of the different STOs of the same l follow a straight line on the logarithmic scale. In quantum chemistry, the even-tempered STOs and GTOs are the most popular choice for constructing systematic and efficient AO basis sets. It can be shown that the overlap between two adjacent STOs or GTOs stays constant, leading to an even coverage of the Hilbert space [119,120]. In this work, we employ four STOs for each l channel, and for simplicity we set $n = l + 1$ and $\beta_l = 2$. Thus, for a given l, the STOs with $\zeta_{l,1}$ and $\zeta_{l,4}$ correspond to the most extended and the most localized functions, respectively. With these choices, the set of STOs is fully specified by the highest angular momentum $l_{\max}$ and the smallest component $\zeta_{l,1} = \alpha_l$ for $l <= l_{\max}$. The parameters α_l are chosen such that the overlaps between the STOs centering on neighboring atoms are vanishingly small, and thus including them in the one-electrons OBS does not cause numerical problems associated with linear dependencies.

In the following, we choose ZnO as the test example to check the possible influence of these highly localized STOs in our NAO framework, since previous experience accumulated in the LAPW community [41,90,117] showed that the localized orbitals have substantial impact on the band gap of ZnO. In Fig. 6 we present the indirect G_0W_0@PBE band gap for
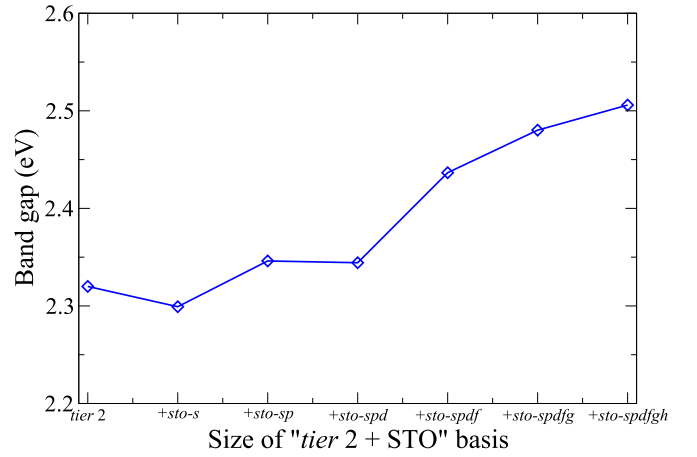

FIG. 6. The variation of the G_0W_0@PBE band gaps for zinc-blende ZnO upon adding STOs of increasing angular momenta to FHI-aims-2009 NAO *tier* 2 basis set. The chosen $\zeta_{l,1}$ values for $l = 0, \ldots, 5$ are 2.5, 5.5, 7.5, 7.5, 10 for Zn and 5, 10.10, 10, 15, 20 for O, respectively. The horizontal dashed line marks the band gap value obtained with *tier* 2 basis set. A combined set of $4f(3.0)5g(3.0)$ hydrogen-like *for_aux* functions and a $8 \times 8 \times 8$ **k** grid were used in the calculations.

ZnO (in its zinc-blende structure) obtained by adding STOs to the *tier* 2 basis set. It can be seen that, as more STOs of increasing angular momenta are included, the obtained band gap gradually increases, evolving from 2.29 eV obtained with the *tier 2* basis set to 2.51 eV obtained with "*tier* 2 + STO-*spdf g h*" basis set (i.e., adding STOs with $l_{\max} = 5$ to *tier* 2). Although the band gap value is not yet saturated for $l_{\max} = 5$, the fully converged band gap is rather unlikely to exceed 2.6 eV, as can be judged from the convergence plot shown in Fig. 6. Thus, this study suggests that complementing the standard NAO basis set with highly localized STOs enlarges the G_0W_0 band gap of ZnO by 0.2 to 0.3 eV. On the one hand, this behavior is in qualitative agreement with what is found in the LAPW framework. On the other hand, the magnitude of the correction brought out by the localized orbitals in the NAO framework is more than a factor of 2 smaller than the LAPW case.

We also performed a similar study as outlined above for Si and MgO. For these systems, adding the highly localized STOs up to $l_{\max} = 5$ gives rise to only a minor increase of the G_0W_0@PBE band gap of just 0.04 eV. In Sec. V we will present benchmark band gap results for a set of crystals computed with both the standard *tier* 2 and the "*tier* 2 + STO-*spdf g h*" basis sets. This will allow us to arrive at a more complete picture of the effects that localized orbitals can bring about in the NAO context. However, adding STOs up to $l_{\max} = 5$ significantly enlarges the size of one-electron basis set. For example, for ZnO the number of orbital basis functions (N_b) per unit cell drastically increases from 91 to 378, and this makes the entire calculation 20 times more expensive.

In a recent paper [108], Zhu and Chan showed that, with relatively small Dunning's cc-pVTZ or even cc-pVDZ GTO basis sets, they can obtain fairly good band gaps for a range of insulating solids, including ZnO. Our preliminary tests with GTOs indicate that the most extended (diffuse) functions in

TABLE I. PBE and G_0W_0@PBE band gaps calculated with FHI-aims, in comparison to the results obtained by FHI-gap [40,109] and the experimental values. The FHI-aims calculations were done with NAO *tier* 2 and $4f(3.0)5g(3.0)$ *for_aux* basis functions. An $8 \times 8 \times 8$ **k** point mesh was used for all crystals except for wurzite structures, whereby a reduced $8 \times 8 \times 5$ **k** grid was used. The WIEN2k PBE results and FHI-gap G_0W_0@PBE results [with both standard LAPW basis ($n_{LO} = 0$) and LAPW+HLOs basis ($n_{LO} = 5$)] are mostly taken from Ref. [90]. The mean deviation (MD) and mean absolute deviation (MAD) for the FHI-aims G_0W_0 results and the LAPW-based ($n_{LO} = 0$) G_0W_0 results are obtained with reference to the LAPW+HLOs ($n_{LO} = 5$) results. The experimental values are directly cited from Ref. [90], originally taken from Refs. [125,126]. Systems with a prefix "WZ" in their names mean the wurzite structure was used in the calculation; otherwise the zinc blende structure was used instead. In all calculations the experimental lattice constants were used. The experimental band gaps are corrected for the ZPR effect, i.e., $E_g^{corr} = E_g(T = 0) - \Delta E_g^{ZPR}$ with the ZPR contribution ΔE_g^{ZPR} given in parentheses. Unless otherwise noted, the ZPR term is estimated from a linear extrapolation of the experimental gap $E_g(T)$ to $T = 0$, as collected in Ref. [127].

Crystals	Exp. (ΔE_g^{ZPR})	FHI-aims			WIEN2K	2cFHI-gap	
		PBE	G_0W_0 (*tier*2)	G_0W_0 (*tier*2+STO)	PBE	G_0W_0 ($n_{LO} = 0$)	G_0W_0 ($n_{LO} = 5$)
AlAs	2.27(−0.039)	1.34(0.10)	2.03	2.23	1.34(0.10)	1.94	2.06
WZ-AlN	6.44(−0.239)	4.21	5.75	5.86	4.14	5.60	5.80
AlP	2.53(−0.023)	1.58	2.32	2.43	1.57	2.25	2.37
AlSb	1.73 (−0.039)	0.95(0.25)	1.47	1.70	1.03(0.22)	1.40	1.50
BAs	1.65(−0.151[a])	1.19	1.85	1.91	1.19	1.78	1.83
BN	6.66(−0.262[b])	4.46	6.32	6.31	4.46	6.04	6.36
BP	2.5,2.2 (−0.106[d])	1.24	1.95	1.99	1.34	2.01	2.11
C	5.85(−0.370)	4.13	5.61	5.60	4.16	5.49	5.69
CdS	–	1.16	1.98	2.06	1.14	1.94	2.06
WZ-CdS	2.64(−0.068)	1.17	1.97	2.05	1.20	2.02	2.19
GaN	3.64(−0.173)	1.61	2.70	2.96	1.68	2.78	3.00
GaP	2.43(−0.085)	1.64	2.15	2.14	1.66	2.05	2.21
LiCl	9.8 (−0.436[d])	6.33	8.55	8.64	6.30	8.56	8.71
LiF	14.48(−0.281[b])	9.20	13.79	13.49	9.28	13.36	14.27
MgO	7.98(−0.154[b])	4.73	7.32	7.36	4.75	7.08	7.52
NaCl	8.6 (−0.098[c])	5.10	7.69	7.80	5.12	7.67	7.92
Si	1.23 (−0.064)	0.61	1.09	1.13	0.56	1.03	1.12
SiC	2.57 (−0.145[c])	1.37	2.38	2.52	1.36	2.23	2.38
ZnO	–	0.69	2.29	2.51	0.70	2.05	2.78
WZ-ZnO	3.60(−0.156)	0.82	2.46	2.70	0.83	2.24	3.01
MD			−0.16	−0.08		−0.28	
MAD			0.17	0.15		0.28	

[a]Theoretical estimates in Ref. [121].
[b]Theoretical estimates in Ref. [122].
[c]Theoretical estimates in Ref. [123].
[d]Theoretical estimates in Ref. [124].

cc-pVDZ/TZ basis sets have to be excluded to circumvent the linear-dependent problem noted above. On the other hand, when this is done, the obtained G_0W_0@PBE band gap for ZnO with such modified cc-pVTZ basis set is on par with the more expensive "$tier2 + $STO-$spdfgh$" basis set. This observation points to the possibility of developing compact and efficient NAO basis sets that are more suitable for all-electron G_0W_0 calculations. More investigations along this line are needed to arrive at more faithful conclusions.

V. BENCHMARK RESULTS

In the previous section, we have examined the convergence behavior of the G_0W_0 band gaps with respect to three numerical factors in our NAO-based implementation. Based on the above convergence tests, it appears that $4f(3.0)5g(3.0)$ *for_aux* functions to generate extra ABFs guarantee a good accuracy under the LRI approximation, and that a $8 \times 8 \times 8$ Γ-inclusive **k** point grid is adequate for the BZ sampling (for insulators) in cubic structures. Regarding the one-electron OBS, the FHI-aims-2009 *tier* 2 basis set seems to be adequate for "simple" systems like Si or MgO, but for systems like ZnO, complementing *tier* 2 with highly localized STOs gives rise to to an increase of the band gap of 0.2 to 0.3 eV. In this section, we will perform benchmark G_0W_0 calculations for a set of semiconductors and insulators. For these calculations, the $4f(3.0)5g(3.0)$ *for_aux* functions are used throughout. As for the **k** grid, an $8 \times 8 \times 8$ mesh is used for all crystals with cubic [zinc blende (ZB) for binary compounds] structure, whereas a reduced $8 \times 8 \times 5$ mesh is used for the wurzite (WZ) structure. For one-electron OBS, both the "*tier* 2" and "*tier* 2 + STO-$spdfg$" basis set will be used. This enables one to assess the overall influence of the highly localized orbitals on the computed G_0W_0 band gaps. Our results calculated in this work will be compared to the reference values obtained using the FHI-gap code [40,109].

In Table I we present our calculated PBE and G_0W_0@PBE band gaps of 21 crystals, covering systems from small gap semiconductors to wide gap insulators, formed with a wide variety of chemical elements. For AlN, CdS, and ZnO,
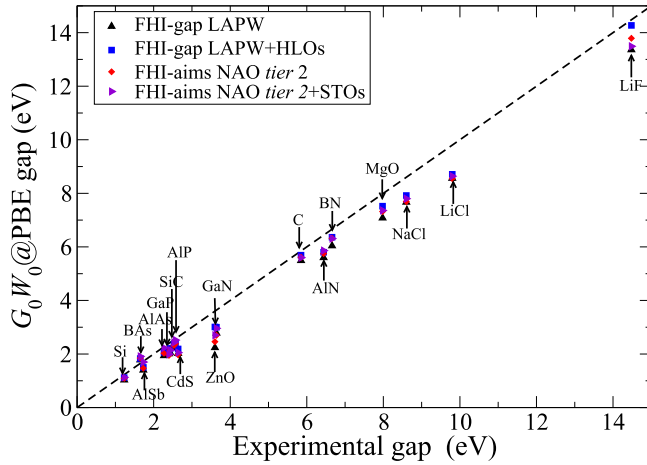
FIG. 7. Calculated G_0W_0 band gaps versus the experimental ones for the materials presented in Table I. For materials where results of both the ZB and WZ structures are calculated, only results for the WZ structure results are presented. The experimental values are corrected for the ZPR effect.

results for both ZB and WZ crystal structures are shown. Also presented are the reference PBE and G_0W_0@PBE band gap results obtained respectively by the WIEN2k [110] and FHI-gap [40,109] codes, as well as the experimental band gap values. The presented experimental values are corrected for the zero-point renormalization (ZPR) effect,

$$E_g^{\text{corr}} = E_g(T=0) - \Delta E_g^{\text{ZPR}}, \tag{41}$$

with $E_g(T=0)$ being the experimental value extrapolated to zero temperature, and ΔE_g^{ZPR} being the ZPR contribution. The calculated G_0W_0@PBE gap values versus the experimental ones are further presented graphically in Fig. 7. The WIEN2k plus FHI-gap results are mostly taken from Ref. [90], except for CdS and LiCl, which are computed for the first time in this work. The FHI-gap calculations were done with $6 \times 6 \times 6$ **k** grid, but tests show that the G_0W_0 band gaps are well converged with their own **k** point convergence strategy [40]. For heavy elements where the spin-orbit coupling (SOC) effect is significant, both the PBE and G_0W_0 band gaps are corrected by a SOC term given in parentheses along with the PBE value. In FHI-aims calculations, the SOC effect is treated using the non-self-consistent second variational method following a self-consistent scalar relativistic calculation [74]. Again the PBE gaps show a remarkable level of agreement between FHI-aims and WIEN2k. The FHI-gap G_0W_0 results were obtained both with the standard LAPW basis set ($n_{\text{LO}} = 0$) and with the newly developed LAPW + HLOs prescription [111,117]. In Ref. [90], Jiang and Blaha demonstrated the influence of the addition of HLOs in the LAPW framework to the *GW* band gaps of a sequence of materials. It was shown that both the G_0W_0 and GW_0 band gaps get enlarged when adding HLOs, but the magnitude of the correction varies from system to system, ranging from less than 0.1 eV to more than 0.7 eV.

From Table I, one may observe that, with few exceptions, the FHI-aims G_0W_0@PBE results obtained with *tier* 2 basis set are found in between the FHI-gap results obtained respectively with the standard LAPW basis set and with the highly

TABLE II. Recently reported G_0W_0 band gaps for WZ-ZnO from various implementations. In the "Framework" column, "AE" means "all-electron" and "NCP" means norm-conserving pseudopotential.

Code	Framework	G_0W_0 gap (eV)		Ref.
		@LDA	@PBE	
FHI-aims	AE + NAO	2.78	2.70	This work
FHI-gap	LAPW	–	3.01	[90]
Jüliech	LAPW	2.99	–	[117]
Exciting	LAPW	2.94	–	[41]
VASP	PAW + PW	2.87	2.76	[32]
Yambo	NCP + PW	2.8	–	[130]
Abinit	NCP + PW	2.8	–	[130]
BerkeleyGW	NCP + PW	2.8	–	[130]
PySCF	AE + GTO	–	3.08	[108]

converged LAPW + HLOs basis set. This behavior can also be clearly seen from the graphical presentation of the data in Fig. 7. Actually, for about half of the materials, the FHI-aims *tier* 2 results are fairly close (within 0.1 eV) to the FHI-gap results with the LAPW + HLOs basis prescription, taken to be the reference here. As indicated in Table I, the mean deviation (MD) and mean absolute deviation (MAD) of the FHI-aims *tier* 2 G_0W_0 results with respect to those obtained from the LAPW + HLOs($n_{\text{LO}} = 5$) basis set (last column in Table I) are substantially smaller than the corresponding MD and MAD values for the standard LAPW G_0W_0 results. However, for systems where the addition of HLOs has a substantial influence, the FHI-aims *tier* 2 results also show an appreciable underestimation of band gaps, as compared to the LAPW + HLOs results. This is particularly true for ZnO, LiF, NaCl, and GaN. A common feature of these materials is that they have a strong ionic character and a direct band gap at the Γ point. For ionic solids, the VBM and CBM originate from different chemical elements (and orbital characters) and hence may evolve differently as one increases the basis size, leading to an overall slower convergence of the band gaps.

In particular, for ZnO, the G_0W_0 calculation is notoriously difficult to converge with respect to the unoccupied states. Since the remarkable finding of Shih *et al.* [128] that the earlier *GW* calculations of ZnO were severely underconverged, different *GW* implementations [32,41,90,117,129,130] have been tested for this system. Although differing in details, most *GW* codes yield a substantial increase of the band gap for ZnO if the calculation is carefully converged with respect to the unoccupied states. As already discussed above, within the LAPW framework, adding the HLOs in the LAPW framework increases the ZnO band gap by more than 0.7 eV [41,90,117]. Regarding the NAO basis set, Table I shows that, for both ZB- and WZ-ZnO, our G_0W_0@PBE gap obtained with the FHI-aims-2009 *tier* 2 basis set are larger by more than 0.2 eV than the corresponding values obtained with the standard LAPW basis set, but are about 0.5 eV smaller than the values obtained with the fully converged LAPW + HLOs basis set. As demonstrated in Sec. IV C 2, within our G_0W_0 implementation, complementing *tier* 2 with highly localized STOs leads to an enlargement of 0.2 to 0.3 eV of the band gap for ZB-ZnO. In Table II we accumulate the G_0W_0 band gap values

for ZnO in wurzite structure from various implementations, as reported in the literature recently. One can see that, despite of the recent efforts, a scatter of 0.2–0.3 eV from different G_0W_0 implementations is still visible.

We then perform G_0W_0@PBE calculations with the "$tier2 + $ STO-$spdfgh$" basis set for all systems and the obtained results are also presented in Table I. Close inspection reveals that the actual impact brought by the localized STOs varies from system to system. For most of the materials, adding the STOs to $tier 2$ leads to an increase of the band gaps, similar to the LAPW + HLOs case. Among these, there are materials where the addition of STOs improves the FHI-aims G_0W_0 band gaps towards the "reference" LAPW + HLOs values. This is in particular true for ZnO, NaCl, GaN, LiCl, and CdS where the improvement is substantial (∼0.1 eV or bigger). However, there are also cases (i.e., AlAs, AlP, AlSb, and SiC) where, after including the STOs, the FHI-aims G_0W_0 band gap values exceed the LAPW + HLOs ones, and the agreement between the two codes *de facto* deteriorates. Finally, we note that in a couple of cases, the addition of STOs leads to essentially no correction (C, BN, GaP) or a negative correction (LiF). Due to the varying impacts of the STOs, the MAD of the FHI-aims G_0W_0 band gaps obtained with the "$tier2 + $ STO-$spdfgh$" basis set with reference to the LAPW + HLOs ones does not show an appreciable improvement, although the MD value gets significantly reduced.

In the *GW* community, obtaining fully converged G_0W_0 results with respect to the one-electron OBS is of high academic interest. In the NAO framework, complementing the standard NAOs with highly localized STOs seems to offer a viable route towards this goal. However, a systematic addition of the STOs to the OBS, as described in Sec. IV C 2, leads to a very large number of one-electron basis functions, and as a consequence, the entire G_0W_0 calculation becomes one order of magnitude more expensive (nine times more expensive in case of NiO). This renders the combination of standard NAO with highly localized STOs not a preferable scheme suitable for routine calculations. Developing computationally more affordable schemes to correct the basis set incompleteness error for G_0W_0 calculations is of current interest [131–134]. In particular, correction schemes that account for the response of the basis functions to the change of the effective potential can significantly improve the level of convergence towards the complete basis set limit [131–133]. This is because one effectively goes beyond the Hilbert space defined by the one-electron basis set by taking into account their variations upon responding to perturbations. For practical purposes, we suggest that the *tier* 2 basis set be used as the default choice for periodic G_0W_0 calculations using FHI-aims. When numerically highly accurate results (band gap values converged within 0.1 eV or better) are needed, one can then check the quality of the obtained G_0W_0 results by further adding STOs to the OBS. Our preliminary investigations indicate that it is highly possible to develop more compact NAO basis sets for performing high-precision G_0W_0 calculations, especially if the responses of basis functions can be efficiently included in the computational scheme.

From this benchmark study, one may also judge that, although the overall agreement of the two all-electron G_0W_0 implementations based on different numerical techniques is

rather encouraging, the discrepancies for certain materials are still too large to be acceptable. The level of agreement seen for the DFT-PBE band structures has not been achieved yet for G_0W_0 calculations. More investigations will be needed to clarify the possible origins.

Finally one may notice that f-electron materials are not included in the test set presented in Table I. Our implementation is all-electron, and the core electron are included explicitly in the calculation. Hence the materials containing f-electrons do not pose formal difficulties. However, the convergence with respect to one-electron NAO basis set, as discussed above, will be even more challenging. We are currently testing our implementation on such materials, and the results will be reported in separate publications. The basis function response correction as mentioned above will be of great help there. In this paper, we described in detail the formulation and algorithms of an all-electron periodic G_0W_0 implementation within the NAO framework. To our knowledge, this is the first all-electron NAO-based G_0W_0 implementation that works with periodic boundary conditions. Our implementation was carried out within the FHI-aims code package [59,67]. With the achievement reported in this work, FHI-aims becomes a code that allows one to carry out both molecular and periodic G_0W_0 calculations, in an all-electron fashion, within a unified numerical framework. We performed systematic convergence tests, and identified a set of computational parameters that can be used as default settings in G_0W_0 calculations to obtain reliable results. With such a default setting, we benchmarked our implementation by computing G_0W_0@PBE band gaps for a set of semiconductors and insulators, and compared the obtained band gaps to the independent Wien2k plus FHI-gap [40,109] results. We found that, with the standard NAO *tier* 2 basis, one can obtain band gaps that are already in fairly good agreement with those obtained by FHI-gap with highly accurate LAPW + HLOs basis set. Challenging situations do exist, like the famous ZnO example, where NAOs suffer from a similar one-electron basis under-convergence issue as other basis frameworks do, though to a somewhat less extent. Complementing the FHI-aims-2009 *tier* 2 NAO basis set with highly localized STOs, one obtains appreciable improvements of the band gap value for a fraction of the materials, including ZnO. However, the computational scheme by complementing NAOs with highly localized STO is extremely expensive and is not suitable for routine calculations. For most practical purposes, we recommend to use the FHI-aims-2009 NAO *tier* 2 basis set, which delivers useful accuracy. Importantly, even this level of G_0W_0 calculations can significantly improve over the computationally cheaper but more empirical hybrid density functional approximations. While hybrid DFT calculations are much more readily convergeable to the complete basis set limit, they cannot be parameterized to fully represent the local electronic structure variations in hybrid materials with conceptually different components (e.g., organic-inorganic interfaces). In contrast, G_0W_0, with practically affordable basis set, can provide a much more uniform level of theory that encompasses the locally relevant screening effects in a natural way.

The compactness and locality of NAOs in principles enable efficient periodic GW implementations, allowing for tackling complex materials at reduced cost. Our benchmark calculations show that high-quality GW results can be obtained using

modest NAO basis sets, except for some challenging systems. Generating compact NAO basis sets that can efficiently represent unoccupied Hilbert space is highly desirable, but requires considerably more efforts that goes beyond the scope of the present work. In this endeavor, independent implementations using systematic basis sets such as LAPW + HLOs will provides invaluable reference results for benchmark purposes.

Our implementation is massively parallel. Due to the local nature of our basis set, our implementation can be readily applied to 1D and 2D systems. Benchmark calculations for the efficiency and scalability of our implementation, as well as its performance for systems with lower dimensions, will be presented in a future paper. Finally we are also working to extend our implementation to treat metallic systems.

VI. CONCLUSION AND OUTLOOK

In this paper, we described in detail the formulation and algorithms of an all-electron periodic G_0W_0 implementation within the NAO framework. To our knowledge, this is the first all-electron NAO-based G_0W_0 implementation that works with periodic boundary conditions. Our implementation was carried out within the FHI-aims code package [59,67]. With the achievement reported in this work, FHI-aims becomes a code that allows one to carry out both molecular and periodic G_0W_0 calculations, in an all-electron fashion, within a unified numerical framework. We performed systematic convergence tests, and identified a set of computational parameters that can be used as default settings in G_0W_0 calculations to obtain reliable results. With such a default setting, we benchmarked our implementation by computing G_0W_0@PBE band gaps for a set of semiconductors and insulators, and compared the obtained band gaps to the independent Wien2k plus FHI-gap [40,109] results. We found that, with the standard NAO *tier* 2 basis, one can obtain band gaps that are already in fairly good agreement with those obtained by FHI-gap with a highly accurate LAPW + HLOs basis set. Challenging situations do exist, like the famous ZnO example, where NAOs suffer from a similar one-electron basis under-convergence issue as other basis frameworks do, though to a somewhat less extent. Complementing the FHI-aims-2009 *tier* 2 NAO basis set with highly localized STOs, one obtains appreciable improvements of the band gap value for a fraction of the materials, including ZnO. However, the computational scheme by complementing NAOs

with highly localized STO is extremely expensive, and is not suitable for routine calculations. For most practical purposes, we recommend using the FHI-aims-2009 NAO *tier* 2 basis set, which delivers useful accuracy. Importantly, even this level of G_0W_0 calculations can significantly improve over the computationally cheaper but more empirical hybrid density functional approximations. While hybrid DFT calculations are much more readily convergeable to the complete basis set limit, they cannot be parameterized to fully represent the local electronic structure variations in hybrid materials with conceptually different components (e.g., organic-inorganic interfaces). In contrast, G_0W_0, with practically affordable basis set, can provide a much more uniform level of theory that encompasses the locally relevant screening effects in a natural way.

The compactness and locality of NAOs in principles enable efficient periodic GW implementations, allowing for tackling complex materials at reduced cost. Our benchmark calculations show that high-quality GW results can be obtained using modest NAO basis sets, except for some challenging systems. Generating compact NAO basis sets that can efficiently represent unoccupied Hilbert space is highly desirable, but requires considerably more efforts that goes beyond the scope of the present work. In this endeavor, independent implementations using systematic basis sets such as LAPW + HLOs will provides invaluable reference results for benchmark purposes.

Our implementation is massively parallel. Due to the local nature of our basis set, our implementation can be readily applied to 1D and 2D systems. Benchmark calculations for the efficiency and scalability of our implementation, as well as its performance for systems with lower dimensions, will be presented in a future paper. Finally we are also working to extend our implementation to treat metallic systems.

ACKNOWLEDGMENTS

[1] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[2] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[3] S. H. Vosko, L. Wilk, and M. Nusair, Can. J. Phys. **58**, 1200 (1980).

[4] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

[5] A. Becke, J. Chem. Phys. **88**, 2547 (1988).

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

[7] A. D. Becke, J. Chem. Phys **98**, 5648 (1993).

[8] J. Heyd, G. E. Scuseria, and M. Ernzerhof, J. Chem. Phys. **118**, 8207 (2003).

[9] T. Shimazaki and Y. Asai, Chem. Phys. Lett. **466**, 91 (2008).

[10] M. A. L. Marques, J. Vidal, M. J. T. Oliveira, L. Reining, and S. Botti, Phys. Rev. B **83**, 035119 (2011).

[11] J. H. Skone, M. Govoni, and G. Galli, Phys. Rev. B **93**, 235106 (2016).

[12] J. He and C. Franchini, J. Phys.: Condens. Matter **29**, 454004 (2017).

[13] Z.-H. Cui, Y.-C. Wang, M.-Y. Zhang, X. Xu, and H. Jiang, J. Phys. Chem. Lett. **9**, 2338 (2018).

[14] A. A. Abrikosov, L. P. Gorkov, and I. E. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics* (Dover Publications, New York, 1975).

[15] A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-Particle Systems* (McGraw-Hill, New York, 1971).

[16] L. Hedin, Phys. Rev. **139**, A796 (1965).

[17] G. Strinati, H. J. Mattausch, and W. Hanke, Phys. Rev. B **25**, 2867 (1982).

[18] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986).

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

[20] F. Aryasetiawan and O. Gunnarsson, Rep. Prog. Phys. **61**, 237 (1998).

[21] W. G. Aulbur, L. Jönsson, and J. W. Wilkins, Solid State Phys.: Adv, Res. Appl. **54**, 1 (2000).

[22] G. Onida, L. Reining, and A. Rubio, Rev. Mod. Phys. **74**, 601 (2002).

[23] P. Rinke, A. Qteish, J. Neugebauer, C. Freysoldt, and M. Scheffler, New J. Phys. **7**, 126 (2005).

[24] N. Marom, F. Caruso, X. Ren, O. T. Hofmann, T. Körzdörfer, J. R. Chelikowsky, A. Rubio, M. Scheffler, and P. Rinke, Phys. Rev. B **86**, 245127 (2012).

[25] D. Golze, M. Dvorak, and P. Rinke, Front. Chem. **7**, 377 (2019).

[26] F. Bruneval and X. Gonze, Phys. Rev. B **78**, 085125 (2008).

[27] J. Deslippe, G. Samsonidze, D. A. Strubbe, M. Jain, M. L.Cohen, and S. G. Louie, Comput. Phys. Commun. **183**, 1269 (2012).

[28] M. Govoni and G. Galli, J. Chem. Theory Comput. **11**, 2680 (2015).

[29] B. Arnaud and M. Alouani, Phys. Rev. B **62**, 4464 (2000).

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

[31] M. Shishkin and G. Kresse, Phys. Rev. B **74**, 035101 (2006).

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

[33] F. Hüser, T. Olsen, and K. S. Thygesen, Phys. Rev. B **87**, 235132 (2013).

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

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

[36] F. Aryasetiawan and O. Gunnarsson, Phys. Rev. Lett. **74**, 3221 (1995).

[37] W. Ku and A. G. Eguiluz, Phys. Rev. Lett. **89**, 126401 (2002).

[38] C. Friedrich, A. Schindlmayr, and S. Blügel, Comp. Phys. Comm. **180**, 347 (2009).

[39] C. Friedrich, S. Blügel, and A. Schindlmayr, Phys. Rev. B **81**, 125102 (2010).

[40] H. Jiang, R. I. Gómez-Abal, X.-Z. Li, C. Meisenbichler, C. Ambrosch-Draxl, and M. Scheffler, Comp. Phys. Commun. **184**, 348 (2013).

[41] D. Nabok, A. Gulans, and C. Draxl, Phys. Rev. B **94**, 035118 (2016).

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

[43] J. Wilhelm and J. Hutter, Phys. Rev. B **95**, 235123 (2017).

[44] J. Wilhelm, D. Golze, L. Talirz, J. Hutter, and C. A. Pignedoli, J. Phys. Chem. Lett. **9**, 306 (2018).

[45] U. von Barth and B. Holm, Phys. Rev. B **54**, 8411 (1996).

[46] M. Shishkin, M. Marsman, and G. Kresse, Phys. Rev. Lett. **99**, 246403 (2007).

[47] S. V. Faleev, M. van Schilfgaarde, and T. Kotani, Phys. Rev. Lett. **93**, 126406 (2004).

[48] M. van Schilfgaarde, T. Kotani, and S. Faleev, Phys. Rev. Lett. **96**, 226402 (2006).

[49] A. Stan, N. E. Dahlen, and R. van Leeuwen, Europhys. Lett. **76**, 298 (2006).

[50] C. Rostgaard, K. W. Jacobsen, and K. S. Thygesen, Phys. Rev. B **81**, 085103 (2010).

[51] F. Caruso, P. Rinke, X. Ren, M. Scheffler, and A. Rubio, Phys. Rev. B **86**, 081102(R) (2012).

[52] F. Caruso, P. Rinke, X. Ren, A. Rubio, and M. Scheffler, Phys. Rev. B **88**, 075105 (2013).

[53] A. L. Kutepov, Phys. Rev. B **94**, 155101 (2016).

[54] A. L. Kutepov, Phys. Rev. B **95**, 195120 (2017).

[55] S.-H. Ke, Phys. Rev. B **84**, 205415 (2011).

[56] X. Blase, C. Attaccalite, and V. Olevano, Phys. Rev. B **83**, 115103 (2011).

[57] D. Foerster, P. Koval, and D. Sánchez-Portal, J. Chem. Phys. **135**, 074105 (2011).

[58] C. Faber, I. Duchemin, T. Deutsch, C. Attaccalite, V. Olevano, and X. Blase, J. Mater. Sci. **47**, 7472 (2012).

[59] X. Ren, P. Rinke, V. Blum, J. Wieferink, A. Tkatchenko, A. Sanfilippo, K. Reuter, and M. Scheffler, New J. Phys. **14**, 053020 (2012).

[60] F. Bruneval and M. A. L. Marques, J. Chem. Theory Comput. **9**, 324 (2012).

[61] M. J. van Setten, F. Weigend, and F. Evers, J. Chem. Theory Comput. **9**, 232 (2013).

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

[63] J. W. Knight, X. Wang, L. Gallandi, O. Dolgounitcheva, X. Ren, J. V. Ortiz, P. Rinke, T. Körzdörfer, and N. Marom, J. Chem. Theory Comput. **12**, 615 (2016).

[64] M. J. van Setten, F. Caruso, S. Sharifzadeh, X. Ren, M. Scheffler, F. Liu, J. Lischner, L. Lin, J. R. Deslippe, S. G. Louie *et al.*, J. Chem. Theory Comput. **11**, 5665 (2015).

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

[66] M. F. Lange and T. C. Berkelbach, J. Chem. Theory Comput. **14**, 4224 (2018).

[67] V. Blum, F. Hanke, R. Gehrke, P. Havu, V. Havu, X. Ren, K. Reuter, and M. Scheffler, Comp. Phys. Comm. **180**, 2175 (2009).

[68] V. Havu, V. Blum, P. Havu, and M. Scheffler, J. Comput. Phys. **228**, 8367 (2009).

[69] A. C. Ihrig, J. Wieferink, I. Y. Zhang, M. Ropo, X. Ren, P. Rinke, M. Scheffler, and V. Blum, New J. Phys. **17**, 093020 (2015).

[70] S. V. Levchenko, X. Ren, J. Wieferink, R. Johanni, P. Rinke, V. Blum, and M. Scheffler, Comp. Phys. Comm. **192**, 60 (2015).

[71] F. Knuth, C. Carbognoa, V. Atalla, V. Bluma, and M. Scheffler, Comp. Phys. Commun. **190**, 33 (2015).

[72] K. Lejaeghere, G. Bihlmayer, T. Björkman, P. Blaha, S. Blügel, V. Blum, D. Caliste, I. E. Castelli, S. J. Clark, A. Dal Corso *et al.*, Science **351**, 1415 (2016).

[73] S. R. Jensen, S. Saha, J. A. Flores-Livas, W. Huhn, V. Blum, S. Goedecker, and L. Frediani, J. Phys. Chem. Lett. **8**, 1449 (2017).

[74] W. P. Huhn and V. Blum, Phys. Rev. Mater. **1**, 033803 (2017).

[75] J. L. Whitten, J. Chem. Phys. **58**, 4496 (1973).

[76] B. I. Dunlap, J. W. D. Connolly, and J. R. Sabin, J. Chem. Phys **71**, 3396 (1979).

[77] M. Feyereisen, G. Fitzgerald, and A. Komornicki, Chem. Phys. Lett. **208**, 359 (1993).

[78] O. Vahtras, J. Almlöf, and M. W. Feyereisen, Chem. Phys. Lett. **213**, 514 (1993).

[79] F. Weigend, M. Häser, H. Patzelt, and R. Ahlrichs, Chem. Phys. Lett. **294**, 143 (1998).

[80] B. Delly, J. Chem. Phys. **113**, 7756 (2000).

[81] K. Koepernik and H. Eschrig, Phys. Rev. B **59**, 1743 (1999).

[82] J. M. Soler, E. Artacho, J. D. Gale, A. García, J. Junquera, P. Ordejón, and D. Sánchez-Portal, J. Phys.: Condens. Matter **14**, 2745 (2002).

[83] T. Ozaki, H. Kino, J. Yu, M. Han, N. Kobayashi, M. Ohfuti, F. Ishii, and T. Ohwaki, OpenMX user's manual, http://www.openmx-square.org (2008).

[84] S. D. Kenney and A. P. Horsfield, Comp. Phys. Commun. **180**, 2616 (2009).

[85] P. Li, X. Liu, M. Chen, P. Lin, X. Ren, L. Lin, C. Yang, and L. He, Comput. Mater. Sci. **112**, 503 (2016).

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

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

[88] D. Golze, J. Wilhelm, M. J. van Setten, and P. Rinke, J. Chem. Theory Comput. **14**, 4856 (2018).

[89] F. Aryasetiawan and O. Gunnarsson, Phys. Rev. B **49**, 16214 (1994).

[90] H. Jiang and P. Blaha, Phys. Rev. B **93**, 115203 (2016).

[91] S. Baroni and R. Resta, Phys. Rev. B **33**, 7017 (1986).

[92] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **35**, 5585 (1987).

[93] P. Merlot, T. Kjrgaard, T. Helgaker, R. Lindh, F. Aquilante, S. Reine, and T. B. Pedersen, J. Comput. Chem. **34**, 1486 (2013).

[94] L. N. Wirz, S. S. Reine, and T. B. Pedersen, J. Chem. Theory Comput. **13**, 4897 (2017).

[95] J. D. Talman, J. Comp. Phys. **29**, 35 (1978).

[96] J. D. Talman, J. Chem. Phys. **80**, 2000 (1984).

[97] J. D. Talman, Int. J. Quant. Chem. **93**, 72 (2003).

[98] A. Kutepov, K. Haule, S. Y. Savrasov, and G. Kotliar, Phys. Rev. B **85**, 155129 (2012).

[99] P. Liu, M. Kaltak, J. Klimeš, and G. Kresse, Phys. Rev. B **94**, 165109 (2016).

[100] C. Liu, J. Kloppenburg, Y. Yao, X. Ren, H. Appel, Y. Kanai, and V. Blum, J. Chem. Phys. **152**, 044105 (2020).

[101] I. Y. Zhang, X. Ren, P. Rinke, V. Blum, and M. Scheffler, New J. Phys. **15**, 123033 (2013).

[102] J. T. H. Dunning, J. Chem. Phys. **90**, 1007 (1989).

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

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

[105] R. Sundararaman and T. A. Arias, Phys. Rev. B **87**, 165122 (2013).

[106] C. Ambrosch-Draxl and J. O. Sofo, Comput. Phys. Commun. **175**, 1 (2006).

[107] S. Massidda, M. Posternak, and A. Baldereschi, Phys. Rev. B **48**, 5058 (1993).

[108] T. Zhu and G. K.-L. Chan, J. Chem. Theory Comput. (2021), doi:10.1021/acs.jctc.0c00704.

[109] R. Gómez-Abal, X. Li, M. Scheffler, and C. Ambrosch-Draxl, Phys. Rev. Lett. **101**, 106404 (2008).

[110] P. Blaha, K. Schwarz, G. K. H. Madsen, D. Kvasnicka, and J. Luitz, WIEN2k, an augmented plane wave + local orbitals program for calculating crystal properties (Tech. Universität Wien, Vienna, 2001).

[111] R. Laskowski and P. Blaha, Phys. Rev. B **89**, 014402 (2014).

[112] M.-Y. Zhang and H. Jiang, Phys. Rev. B **100**, 205123 (2019).

[113] H. Jiang, Phys. Rev. B **97**, 245132 (2018).

[114] F. Weigend and R. Ahlrichs, Phys. Chem. Chem. Phys. **7**, 3297 (2005).

[115] T. Helgaker, W. K. H. Koch, and J. Noga, J. Chem. Phys. **106**, 9639 (1997).

[116] I. Y. Zhang, A. J. Logsdail, X. Ren, S. V. Levchenko, L. Ghiringhelli, and M. Scheffler, New J. Phys. **21**, 013025 (2019).

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

[118] R. C. Raffenetti, J. Chem. Phys. **59**, 5936 (1973).

[119] C. M. Reeves and M. C. Harrison, J. Chem. Phys. **39**, 11 (1963).

[120] I. Cherkes, S. Klaiman, and N. Moiseyev, Int. J. Quan. Chem. **109**, 2996 (2009).

[121] I. Bravić and B. Monserrat, Phys. Rev. Mater. **3**, 065402 (2019).

[122] G. Antonius, S. Poncé, E. Lantagne-Hurtubise, G. Auclair, X. Gonze, and M. Côté, Phys. Rev. B **92**, 085137 (2015).

[123] Y. Zhang, Z. Wang, J. Xi, and J. Yang, J. Phys.: Condens. Matter **32**, 475503 (2020).

[124] Honghui Shang, Jin Zhao, and Jinlong Yang (unpublished).

[125] T. Chiang, K. Frank, H. Freund, A. Goldmann, F. J. Himpsel, U. Karlsson, R. Leckey, and W. Schneider, *Landolt-Börnstein, New Series, III/23a: Electronic Structure of Solids: Photoemission Spectra and Related Data* (Springer, Berlin, 1989).

[126] O. Madelung, *Semiconductors: Data Handbook*, 3rd ed. (Springer-Verlag, New York, 2004).

[127] M. Cardona and M. L. W. Thewalt, Rev. Mod. Phys. **77**, 1173 (2005).

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

[129] 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(R) (2011).

[130] T. Rangel, M. Del Ben, D. Varsano, G. Antonius, F. Bruneval, F. H. da Jornada, M. J.van Setten, O. K. Orhan, D. D. O'Regan, A. Canning *et al.*, Comput. Phys. Commun. **255**, 107242 (2020).

[131] M. Betzinger, C. Friedrich, A. Görling, and S. Blügel, Phys. Rev. B **85**, 245124 (2012).

[132] M. Betzinger, C. Friedrich, and S. Blügel, Phys. Rev. B **88**, 075130 (2013).

[133] M. Betzinger, C. Friedrich, A. Görling, and S. Blügel, Phys. Rev. B **92**, 245101 (2015).

[134] P.-F. Loos, B. Pradines, A. Scemama, E. Giner, and J. Toulouse, J. Chem. Theory Comput. **16**, 1018 (2020).