


Maximally localized Wannier functions within the (L)APW + LO methodSebastian Tillack ^{*}, Andris Gulans, and Claudia Draxl*Institut für Physik and IRIS Adlershof, Humboldt-Universität zu Berlin, Berlin, Germany*

(Received 2 September 2019; accepted 15 May 2020; published 1 June 2020)

We present a robust algorithm that computes (maximally localized) Wannier functions without the need of providing an initial guess. Instead, a suitable starting point is constructed automatically from so-called local orbitals which are fundamental building blocks of the basis set within (linearized) augmented plane-wave methods. Our approach is applied to a vast variety of materials such as (semi)metals, bulk and low-dimensional semiconductors, and complex inorganic-organic hybrid interfaces. For the interpolation of electronic single-particle energies, an accuracy in the meV range can be easily achieved. We exemplify the capabilities of our method by the calculation of the joint density of states in aluminum, (generalized) Kohn-Sham and quasiparticle band structures in various semiconductors, and the electronic structure of β -Ga₂O₃, including electron and hole effective masses.

DOI: [10.1103/PhysRevB.101.235102](https://doi.org/10.1103/PhysRevB.101.235102)**I. INTRODUCTION**

In the past two decades, maximally localized Wannier functions (MLWFs) became a well-established tool in solid-state calculations. Due to their localized nature they are superior to the equivalent Bloch representation in terms of chemical interpretation. They provide inexpensive access to both single-particle eigenvalues and eigenfunctions at any point in reciprocal space in terms of the so-called Wannier interpolation scheme. Typically, the starting point for the calculation of MLWFs for a quantum mechanical system is a set of single-particle Kohn-Sham (KS) wave functions obtained from density-functional theory (DFT). The nowadays most commonly used approach to MLWFs in solids is based on works by Marzari and Vanderbilt [1] (MV) and Souza and co-workers [2]. Given a set of single-particle orbitals, the MV algorithm approaches a set of MLWFs by an iterative minimization of the target functional Ω , measuring the spread of the Wannier functions (WFs). In general, this optimization problem is nonlinear and high dimensional. Consequently, the result may strongly depend on the starting point for the minimization, and the algorithm can be easily trapped in false local minima unless a sufficiently good starting point is provided. The latter is usually done by specifying a set of projection functions that approximate the sought WFs. In many cases, however, it is not easy to find a reasonable guess for the projection functions. This is particularly difficult when it comes to the construction of WFs from wide energy ranges of entangled bands, in systems with complex geometries or when the states are strongly hybridized. Although a recent study has proposed methods that are not based on projection [3], the MV algorithm is still the standard approach in the construction of MLWFs. A great improvement of the projection method has been made by Mustafa and co-workers [4] who have introduced an algorithm that automatically constructs

a set of optimized projection functions (OPFs) from a large pool of localized trial orbitals. For the valence bands of many materials such as SiO₂ and Cr₂O₃, the spread Ω of the initial guess obtained from the OPF method was shown to be just a few percent larger than the aimed global minimum [4].

Among the various ways of solving the KS equations of DFT, the full-potential linearized augmented plane-wave (LAPW) method, is considered to be the most accurate one. Highest numerical precision can be reached by selectively adding so-called local orbitals (LOs) to the LAPW basis set. The LOs are strongly localized atomiclike functions. Hence, it is natural to employ the LOs in the construction of WFs within the LAPW + LO method. In this work, we combine the well-established MV approach [1,2] with the more recent OPF technique [4]. We employ the latter to automatically construct suitable initial guesses to MLWFs from a set of LOs. We demonstrate that this approach is capable to robustly construct MLWFs in a vast variety of materials without the need of manually providing a starting point. We apply the obtained WFs to study chemical bonding in a series of elemental and binary semiconductors. Further, electronic properties are calculated for more complex bulk and two-dimensional semiconductors as well as a hybrid organic-inorganic interface by the use of WFs based on (generalized) KS states and quasiparticle energies, respectively. We demonstrate that Wannier interpolation is capable to easily provide electronic energies with an accuracy in the meV range over the entire Brillouin zone.

II. METHODOLOGY**A. Theory of Wannier functions**

Here, we briefly discuss the basic steps in the construction of MLWFs and their application to interpolation. For an extensive overview over the MV approach, we refer to Ref. [5].

Let $\psi_{n,\mathbf{k}}(\mathbf{r})$ be a set of single-particle Bloch wave functions describing a quantum-mechanical system as they may be obtained from a DFT calculation or any other method providing

^{*}sebastian.tillack@physik.hu-berlin.de

single-(quasi)particle eigenstates. In solids, the description of a quantum state in terms of Bloch functions is the natural choice, and the quantum numbers n and \mathbf{k} label an energy band and a wave vector in the first Brillouin zone (BZ), respectively. The Bloch formalism, however, is not the only way to describe quantum states in solids, and WFs provide an alternative representation. The transformation between a Bloch function $\phi_{n,\mathbf{k}}$ and a WF $w_{n,\mathbf{R}}$ reads as [6]

$$w_{n,\mathbf{R}}(\mathbf{r}) = \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{R}} \phi_{n,\mathbf{k}}(\mathbf{r}), \quad (1)$$

where \mathbf{R} is a real-space lattice vector labeling a unit cell within a supercell conjugate to the \mathbf{k} -point grid. Equation (1) holds for Bloch functions describing an isolated energy band. In solids, typically only deep-lying (semi)core states form isolated bands. Therefore, it is desirable to generalize this transformation to a multitude of bands. To this end, we first consider an isolated group of energy bands, i.e., a group of J bands that remains separated from all other bands by a finite energy gap throughout the BZ. The states $\psi_{n,\mathbf{k}}$ within this group span a subspace of the full space of solutions to the single-particle problem. Thus, they can be mixed according to some unitary transformation $U^{\mathbf{k}}$. The mixed states

$$\phi_{n,\mathbf{k}}(\mathbf{r}) = \sum_{m=1}^J U_{mn}^{\mathbf{k}} \psi_{m,\mathbf{k}}(\mathbf{r}) \quad (2)$$

form an equally valid basis of the considered subspace and so do the WFs constructed according to Eq. (1). The unitary $J \times J$ matrix $U^{\mathbf{k}}$ reflects a generalization of the phase freedom of a single state and can be chosen freely. This freedom allows for the construction of WFs that are maximally localized according to some localization criterion. From another perspective, the matrices $U^{\mathbf{k}}$ define a gauge and are chosen such that the mixed states $\phi_{n,\mathbf{k}}$ are as smooth in \mathbf{k} as possible, and consequently the Fourier transform in Eq. (1) results in spatially well-localized WFs. Although the valence bands in insulating or semiconducting materials usually form such isolated groups, the conduction bands or the bands in metals often do not. In the case of such *entangled* bands, first, at each \mathbf{k} point a J -dimensional subspace

$$\tilde{\psi}_{m,\mathbf{k}}(\mathbf{r}) = \sum_{\mu=1}^{\mathcal{J}_{\mathbf{k}}} \mathcal{U}_{\mu m}^{\mathbf{k}} \psi_{\mu,\mathbf{k}}(\mathbf{r}) \quad (3)$$

has to be disentangled from the $\mathcal{J}_{\mathbf{k}} \geq J$ bands that fall inside a given (outer) energy window. This subspace is described by a rectangular $\mathcal{J}_{\mathbf{k}} \times J$ matrix $\mathcal{U}^{\mathbf{k}}$ which is semiunitary (i.e., $\mathcal{U}^{\mathbf{k}\dagger} \mathcal{U}^{\mathbf{k}} = \mathbb{1}_J$). Here, J is the number of WFs one aims to construct from the bands inside an energy window of interest, and the $\psi_{\mu,\mathbf{k}}$ are single-particle wave functions whose eigenvalues fall inside that window. Furthermore, a second (inner) energy window can be introduced within which the states $\tilde{\psi}_{m,\mathbf{k}}$ in the disentangled subspace remain unchanged (i.e., $\mathcal{U}_{\mu m}^{\mathbf{k}} = \delta_{\mu m}$ for all states μ, m inside the inner window). Once the J -dimensional subspace is found, the construction of MLWFs is equivalent to the case of isolated bands with $\psi_{m,\mathbf{k}}$ replaced by $\tilde{\psi}_{m,\mathbf{k}}$ in Eq. (2).

The MLWFs obtained from the above procedure form an excellent tight-binding basis which makes them suitable for

an effective reciprocal-space interpolation in terms of a Slater-Koster interpolation [7]. This Wannier interpolation scheme is based on the inversion of Eq. (1) at an arbitrary point \mathbf{q} in reciprocal space for which an interpolation is needed:

$$\phi_{m,\mathbf{q}}(\mathbf{r}) = \sum_{\mathbf{R}} e^{i\mathbf{q}\cdot\mathbf{R}} w_{m,0}(\mathbf{r} - \mathbf{R}). \quad (4)$$

Equation (4) describes the classical tight-binding approach and diagonalizing the Hamiltonian matrix

$$\mathcal{H}_{mn}^{\mathbf{q}} = \langle \phi_{m,\mathbf{q}} | \hat{\mathbf{H}} | \phi_{n,\mathbf{q}} \rangle \quad (5)$$

gives rise to the single-particle eigenvalues $\epsilon_n^{\mathbf{q}}$ and eigenfunctions $\psi_{n,\mathbf{q}}$ at \mathbf{q} expressed in terms of the auxiliary basis $\phi_{m,\mathbf{q}}$. The reason for the efficiency of this approach is that $\mathcal{H}^{\mathbf{q}}$ is easy to construct and typically much smaller than the Hamiltonian expanded in the original first-principles basis in which the states $\psi_{n,\mathbf{k}}$ are expressed. $\mathcal{H}^{\mathbf{q}}$ has the dimension J (the number of bands under consideration) and therefore is easily diagonalized using standard linear-algebra routines.

B. (L)APW + LO method

The approach described in detail below has been implemented into the full-potential all-electron code exciting [8] which is a realization of the (L)APW + LO method. This package implements DFT and many-body perturbation theory (MBPT). The latter is used to compute quasiparticle energies within the G_0W_0 approximation.

The APW method employs a partitioning of the unit cell into so-called muffin-tin spheres (nonoverlapping spheres centered at the nuclei) and an interstitial region (space between the muffin-tin spheres). The basis functions are plane waves in the interstitial region which are smoothly augmented into the muffin-tin spheres by atomiclike functions. The latter are expanded in terms of spherical harmonics around the nuclei. The corresponding radial functions $u_l(r; E_l)$ are solutions of the radial Schrödinger equation and parametrically depend on the energy E_l . In principle, the parameters E_l have to be set to the band energies. In practice, however, these are not known *a priori*, and the basis itself would depend on the solution of the KS equations resulting in a nonlinear eigenvalue equation. In order to linearize the eigenvalue problem, E_l is set to a fixed value typically chosen to lie inside the respective band. In order to add more variational flexibility, the energy derivatives $\dot{u}(r; E_l) = \partial u(r; E_l) / \partial E_l$ can be added to the radial functions resulting in LAPWs.

This basis set can be further extended by the addition of so-called local orbitals (LOs). These functions are nonvanishing only inside one particular muffin-tin sphere at the atomic site \mathbf{R}_{α_L} , where they are given by

$$\phi_L(\mathbf{r}) = \left[\sum_{\xi} a_{\xi}^L u_{l_L, \xi}^{\alpha_L}(|\mathbf{r}_{\alpha_L}|; E_{l_L}) \right] Y_{l_L m_L}(\widehat{\mathbf{r}_{\alpha_L}}), \quad (6)$$

with $\mathbf{r}_{\alpha_L} = \mathbf{r} - \mathbf{R}_{\alpha_L}$. The coefficients a_{ξ}^L are chosen such that ϕ_L is normalized and continuous at the muffin-tin boundary. The radial functions $u_{l_L, \xi}^{\alpha_L}(r; E)$ are solutions of the radial Schrödinger equation with a spherically symmetric potential inside the muffin-tin sphere, and the parameter ξ denotes the linearization order (order of the derivative with respect to

the energy parameter E_l). The addition of LOs results in a highly flexible and tunable basis set and allows for a smaller plane-wave cutoff.

Whenever high numerical precision is demanded, the full-potential (L)APW + LO method is considered the gold standard approach to first-principles calculations based on DFT and allows for the most precise numerical treatment of both ground-state [9] and excited-state properties [10].

C. Wannier functions from local orbitals

The MV approach aims to find a set of unitary matrices $U^{\mathbf{k}}$ that minimizes the WF spread

$$\Omega = \sum_n [\langle w_{n,0} | r^2 | w_{n,0} \rangle - \langle w_{n,0} | \mathbf{r} | w_{n,0} \rangle^2]. \quad (7)$$

In order to ensure a convergence of the iterative minimization of Ω and to minimize the risk of becoming trapped in false local minima, a good starting point is indispensable. In our implementation, we avoid to manually provide suitable projection functions by the use of the OPF method [4]. We find J suitable projection functions as a linear combination of $M \geq J$ localized trial orbitals $h_i(\mathbf{r})$ such that they minimize an approximate spread Ω [see Eq. (40) of Ref. [4]]. Then, this guess is taken as the starting point for the MV approach. In the original work of Mustafa and co-workers [4], the OPF method was introduced to construct MLWFs for isolated groups of bands. This, however, is no conceptual limitation and we demonstrate in this work that it can also be successfully applied to entangled states. The OPF technique is capable of finding WFs that are already very close to the result of the MV approach for isolated valence bands. The guess from OPFs usually is not as good for conduction bands, but our goal is to find a suitable initial guess for the MV approach rather than MLWFs directly. We show that the construction of OPFs can be automatized, which is particularly useful for systems where it is difficult to use physical intuition for constructing suitable projection functions. We construct OPFs from a pool of LOs. The choice of LOs as trial functions is appealing for several reasons: (i) They are already well localized by definition (nonzero only inside one muffin-tin sphere). (ii) They fit any specific problem at hand since they depend on the actual potential in the system. (iii) All integrals needed are already available within the (L)APW + LO method. In practice, we proceed as follows. For each atom α , we construct LOs in the form

$$\begin{aligned} \phi_L(\mathbf{r}) = & [a_0 u_{nl,0}(|\mathbf{r}_\alpha|; E_{nl}) \\ & + a_\xi u_{nl,\xi}(|\mathbf{r}_\alpha|; E_{nl})] Y_{lm}(\hat{\mathbf{r}}_\alpha). \end{aligned} \quad (8)$$

The principal quantum number n and the angular quantum number l are varied according to the Aufbau principle to obtain a specified number of projection functions per atom. The order of energy derivatives ξ is limited to 1 or 2. If LOs with $\xi > 1$ are included, it may happen that some of them become linearly dependent. In this case, we remove linearly dependent LOs from the pool. A set built this way may contain LOs that are not included in the (L)APW + LO basis set in which the self-consistent potential was computed. As discussed in Ref. [4], it may be advantageous to also include projection functions centered at atoms in neighboring unit

cells to cover all possible bonds. In all examples presented in this work, we find that this is not necessary and all calculations are performed including projectors in the ‘‘home’’ unit cell only. For the simple compounds presented in Sec. III A, the inclusion of neighboring unit cells leads to the equivalent set of WFs, possibly just centered at another equivalent bond.

The OPF method provides a guess to the searched MLWFs. In order to fulfill the orthonormality constraint of WFs, also the set of localized trial orbitals h_i needs to be orthonormal. We construct such a set by finding those linear combinations

$$h_i(\mathbf{r}) = \sum_L C_{Li} \phi_L(\mathbf{r}) \quad (9)$$

that maximize the projection onto the considered J -dimensional (disentangled) subspace. To this end, we minimize the functional

$$\mathcal{P} = J - \sum_i \langle h_i | \hat{\mathbf{P}} | h_i \rangle + \sum_{i,j} \Lambda_{ij} [\langle h_i | h_j \rangle - \delta_{ij}], \quad (10)$$

where $\hat{\mathbf{P}} = 1/N_{\mathbf{k}} \sum_{n,\mathbf{k}} |\Psi_{n,\mathbf{k}} \rangle \langle \Psi_{n,\mathbf{k}}|$, and the second term describes the orthonormality constraint with Lagrangian multipliers Λ_{ij} . With this definition, $\mathcal{P} = 0$ would correspond to the case where the set of LOs fully spans the considered subspace (which in practice will not be reached since LOs are strictly zero in the interstitial region). Using Eq. (9), the above-described optimization problem is equivalent to solving the generalized eigenproblem $PC = SC\Lambda$ with $P_{LL'} = \langle \phi_L | \hat{\mathbf{P}} | \phi_{L'} \rangle$, $S_{LL'} = \langle \phi_L | \phi_{L'} \rangle$, and $\Lambda_{LL'} = \lambda_L \delta_{LL'}$. The eigenvalues are positive, and $\mathcal{P} = J - \sum_L \lambda_L \geq 0$. We use only the eigenvectors C_i associated to eigenvalues λ_i that are larger than a specified threshold in order to reduce the total number of trial orbitals h_i and thus the computational cost of constructing OPFs, which usually amounts for only a small fraction of the entire cost for constructing the WFs. The most demanding task is still the computation of the overlaps $M^{\mathbf{k},\mathbf{b}} = \langle u_{\mathbf{k}} | u_{\mathbf{k}+\mathbf{b}} \rangle$ (see, e.g., Ref. [1] for further details) and the final optimization. The time needed to find the OPFs is usually one to two orders of magnitude smaller. As a rule of thumb, the entire cost for finding well-localized WFs is small compared to a DFT calculation with (semi)local exchange-correlation functional and negligible compared to hybrid or *GW* calculations.

III. RESULTS

A. Construction and chemical analysis

The localized nature of WFs and their formal exactness make the Wannier representation superior to the Bloch representation in terms of interpretation and chemical analysis. As an example for the chemical interpretation of MLWFs, we consider various group IV and III–V compounds crystallizing in the diamond or zinc-blende structure. All 16 considered materials (listed in Table I) are semiconductors and exhibit similar electronic properties. In particular, they form an isolated group of four distinct valence bands with hybridized sp^3 character for which we construct a set of four MLWFs. They transform into one another under symmetry operations, and each of them corresponds to one of the four tetrahedral bonds that each atom in these systems forms. The results

TABLE I. WF spreads Ω and shifts σ of the WF centers calculated for the valence bands in 16 group IV and III–V compounds in the diamond (D) and zinc-blende (ZB) structures. The given lattice constants a are adopted from Ref. [11].

		a (Å)	Ω (Å ²)		σ	
			This work	Ref. [11]	This work	Ref. [11]
Si	D	5.431	8.200	8.232	0.000	0.000
Ge	D	5.658	10.078	10.116	0.000	0.000
Sn	D	6.490	13.752	13.801	0.000	0.000
BP	ZB	4.540	5.532	5.479	0.034	0.032
BA _s	ZB	4.777	6.207	6.211	0.048	0.052
GaSb	ZB	6.100	11.390	11.527	0.146	0.154
InSb	ZB	6.480	12.484	12.251	0.202	0.220
GaP	ZB	5.450	8.071	7.637	0.220	0.240
GaAs	ZB	5.650	9.266	8.871	0.222	0.236
AlSb	ZB	6.140	10.275	10.135	0.234	0.228
InP	ZB	5.870	9.370	8.492	0.274	0.308
InAs	ZB	6.060	10.730	10.138	0.274	0.302
SiC	ZB	4.360	4.741	4.651	0.302	0.308
AlAs	ZB	5.660	8.197	8.090	0.310	0.310
AIP	ZB	5.460	7.250	7.146	0.312	0.314
BN	ZB	3.620	2.857	2.820	0.314	0.316

are depicted in Fig. 1 for the bonding states. Indeed, the corresponding WFs have the character of a bonding σ orbital, i.e., they are formed by a linear combination of the two sp^3 -hybridized orbitals from both bonding atoms [1]. From visual inspection of these orbitals, qualitative information about the bond character can be gained. For purely covalently bound

systems (e.g., Ge) the WFs are symmetric and centered right in the middle of the bond while for more ionic bonds (e.g., c-BN) they are asymmetric and pushed toward the more electronegative atom (nitrogen in this example). Built upon this observation, Abu-Farsakh and Qteish [11] proposed a first-principles parameter-free ionicity scale based on the position of the WF centers $\langle \mathbf{r} \rangle_n = \langle w_{n,0} | \mathbf{r} | w_{n,0} \rangle$. For 32 compounds of the type $A^N B^{8-N}$ ($N = 1, \dots, 4$), they defined the bond ionicity based on the parameter σ , describing the shift of the WF center away from the bond center ($\sigma = 0$) toward the anion ($\sigma = 1$). We use their findings to check our automated construction of MLWFs against an existing implementation for the 16 compounds studied here. As far as possible, the numerical parameters (lattice constants, \mathbf{k} grids for obtaining the density and WFs, xc type) are adopted from Ref. [11]. The results are shown in Fig. 2. In all examples, our implementation finds the global minimum of the spread Ω . It is worth noting that for this class of materials with bond-centered WFs the choice of LOs (which are strictly atom centered and even vanish along the bond direction) as projection functions seems counterintuitive. Indeed, we find that the use of mere s - and p -like LOs as projection functions can result in a local minimum of the spread Ω corresponding to atom-centered WFs. However, this issue is fully resolved by employing suitable linear combinations of LOs obtained by the use of the OPF method. In Table I we present both the spread Ω and the shift σ for all 16 materials. With a relative deviation Δ of at most 10%, both quantities are in good agreement with Ref. [11] (top of Fig 2). We attribute these discrepancies to different approximations in the underlying first-principles calculation resulting in different densities and wave functions.

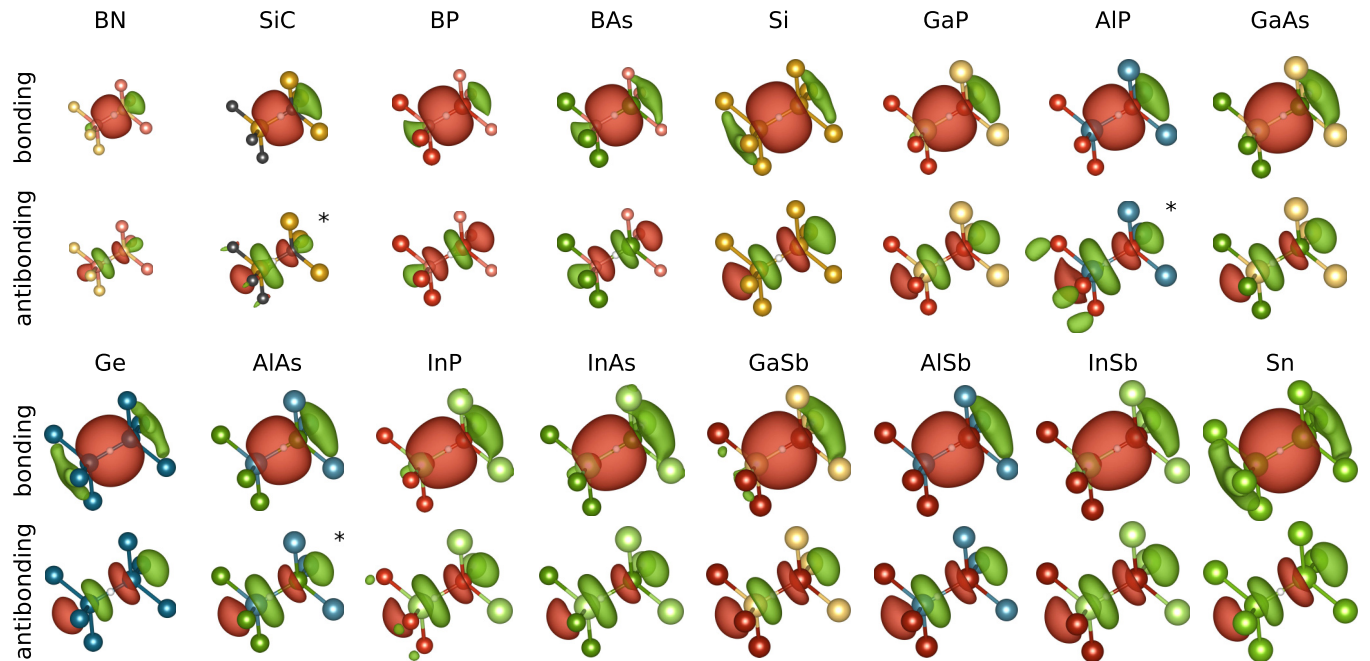


FIG. 1. MLWFs corresponding to the (bonding) valence and lowest (antibonding) conduction bands in 16 group IV and III–IV compounds in the diamond and zinc-blende structures. The white sphere on the central-bond axes illustrates the WF center. All functions are real valued, and surfaces for the same positive (negative) isovalue are shown in red (green). The isovalue is chosen such that the squared modulus integrated over the enclosed volume is 0.9 (0.5) for the bonding (antibonding) WFs. The antibonding orbitals marked with an asterisk are linear combinations of the four corresponding WFs (see also Fig. 3).

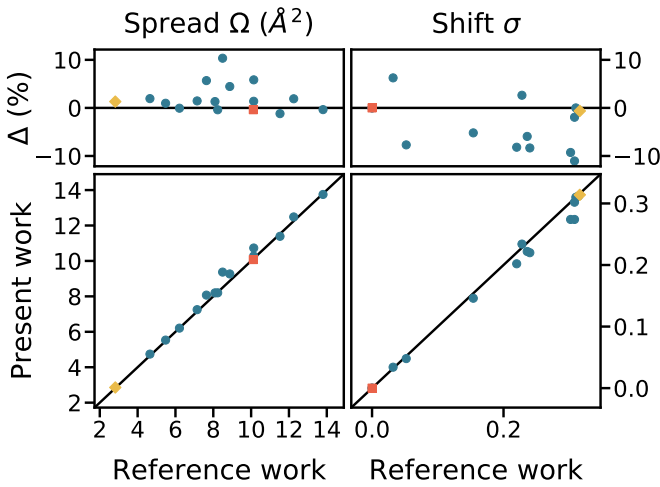


FIG. 2. Spread Ω of the WFs (left) and shift σ of their centers (right) for 16 group IV and III–V compounds. The results obtained within this work are compared against Ref. [11]. For both quantities, the relative deviation $\Delta = (A - A_{\text{ref}})/A_{\text{ref}}$ ($A = \Omega, \sigma$) is within 10% (top). One purely covalent system (Ge) and the most ionic compound (c-BN) are highlighted by a red square and a yellow diamond, respectively.

While we employ a full-potential all-electron approach within the (L)APW + LO basis, in Ref. [11] pseudopotentials and plane waves were used.

In contrast to the valence bands, the conduction bands do not form isolated groups, and the disentanglement procedure needs to be employed. For all 16 compounds, we disentangle four WFs from the lowest conduction bands. Due to the differences in the band dispersion among the materials, there is no single meaningful energy window that would fit all systems. Thus, the energy windows need to be selected individually. In all cases, the lower boundary of the window is set to the conduction band minimum (CBM). The upper boundary is set to the lowest possible energy such that at each \mathbf{k} point there are at least four bands contained in the window. Similar as for the valence bands, the algorithm converges to a set of four equivalent bond-centered WFs for each material. The character of the corresponding WFs is now that of an antibonding σ^* orbital (see Fig. 1) which is in accordance with the findings for Si in Ref. [2]. In Table II we compare both the total spread Ω and its gauge-independent component Ω_1 as well as the position of the WF centers for the valence and the lowest conduction bands. As expected, the spread Ω for the conduction bands is considerably larger than for the valence bands. Further, the fraction of the total spread that comes from the gauge-independent part is smaller for the conduction bands which is mainly due to a relative increase in the off-diagonal term Ω_{OD} (see, e.g., Ref. [1] for the decomposition of the total spread Ω). For the elemental group IV materials, the antibonding WFs are symmetric and centered right in the middle of the bond. For the binary compounds, however, we notice a change in the position of the WF centers. While for the valence bands it is located closer to the more electronegative atom, the opposite behavior is found for the conduction bands in most materials (expressed by negative

TABLE II. WF spreads Ω and shifts σ of the WF centers calculated for the valence and conduction bands in 16 group IV and III–V compounds.

	Valence			Conduction		
	Ω (\AA^2)	Ω_1 (\AA^2)	σ	Ω (\AA^2)	Ω_1 (\AA^2)	σ
Si	8.200	7.696	0.000	29.565	25.893	0.000
Ge	10.078	9.527	0.000	29.333	26.395	0.000
Sn	13.752	12.963	0.000	38.297	34.488	0.000
BP	5.532	5.023	0.034	19.011	16.592	0.290
BAs	6.207	5.742	0.048	19.553	17.294	0.159
GaSb	11.390	10.721	0.146	33.812	30.257	−0.208
InSb	12.484	11.836	0.202	37.727	33.490	−0.249
GaP	8.071	7.533	0.220	28.081	24.215	−0.201
GaAs	9.266	8.687	0.222	28.754	25.424	−0.276
AlSb	10.275	9.593	0.234	34.635	30.492	−0.347
InP	9.370	8.731	0.274	30.056	26.076	−0.350
InAs	10.730	10.010	0.274	30.868	27.084	−0.381
SiC	4.741	4.140	0.302	16.067	13.438	−0.203
AlAs	8.197	7.549	0.310	29.111	25.236	−0.352
AIP	7.250	6.557	0.312	30.920	23.793	−0.272
BN	2.857	2.481	0.314	9.113	7.572	−0.245

values for the shift σ). The magnitude of σ , however, is not anymore directly related to the ionicity based on the shift of the valence band WFs, i.e., a material that is more polar than another one may show a σ that is smaller in magnitude. In our calculations, SiC, AlAs, and AIP stand out from the rest. For these highly polar materials we obtain a set of four WFs which differ from the σ^* -type orbitals but turn out to be linear combinations of them instead (see Fig. 3).

Disentanglement is required also in another scenario, namely, the case of metals and semimetals. First, we consider single-layer graphene. We carry out a DFT calculation within the local-density approximation (LDA) using a $21 \times 21 \times 1$ \mathbf{k} mesh and disentangle a set of five WFs. They correspond to the three occupied and the two partially occupied bands. These bands are fixed inside the inner window up to the onset of the parabolic free-electron bands. As in Ref. [5], we obtain a set of three bonding σ -like WFs corresponding to the occupied bands and two atom-centered p_z -like WFs corresponding to the partially occupied bands. While in Ref. [5] the initial guess was obtained from projection onto atom-centered p_z and sp^2 orbitals (based on physical intuition), no such manual

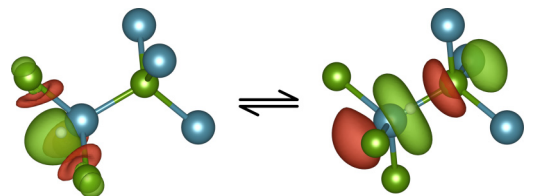


FIG. 3. One of four symmetry-equivalent WFs corresponding to the four lowest conduction bands in AlAs (left) and one of four corresponding antibonding σ^* orbitals (right). Both sets can be expressed as linear combinations of one another. The same behavior is found in SiC and AIP.

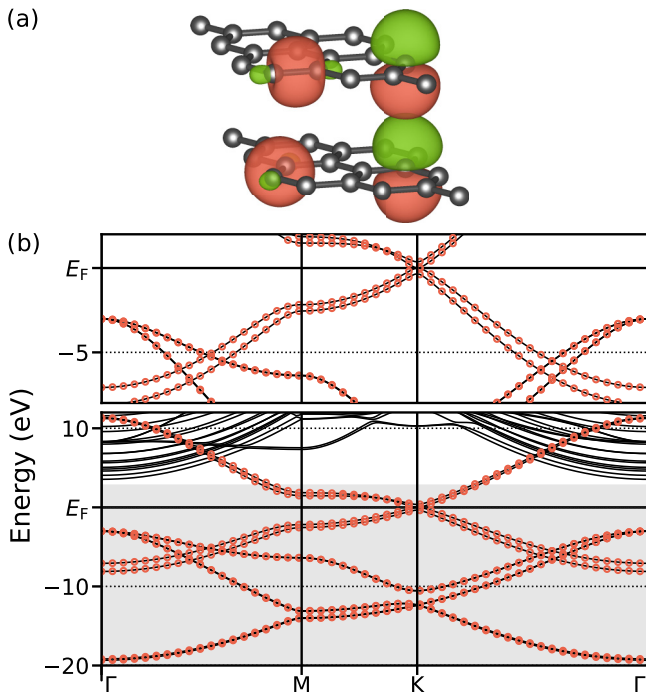


FIG. 4. (a) Two of six σ -like WFs (left orbitals) and two of four p_z -like WFs (right orbitals) in two-layer graphene. The isovalue is chosen such that the squared modulus of the WFs integrated over the enclosed volume is 0.9. (b) DFT band structure (solid line) and Wannier interpolated bands (red circles). The outer window for the calculation of WFs is set to the energy range displayed in the lower panel. The gray shaded area represents the inner energy window. The upper panel magnifies the highest 10 eV of the inner window.

input is required in the present implementation. Our initial guess from OPFs obtained from eight trial orbitals according to Eq. (9) yields a total spread of 3.92 \AA^2 which is already very close to the total minimum of 3.88 \AA^2 obtained from a subsequent minimization. The individual WFs have a spread of 0.62 and 1.01 \AA^2 for the σ and p_z orbitals, respectively. The same set of WFs would be obtained by treating the fully and partially occupied bands separately. We repeat the same procedure for two- and three-layer graphene in AB and ABA stacking, respectively, with an interlayer distance of 3.35 \AA . As a result, we obtain an equivalent set of three σ and two p_z orbitals for each additional layer. Again, the initial guess of the spread obtained from OPFs (given in parentheses) is very close to the total minimum of 7.81 \AA^2 (7.83 \AA^2) and 11.74 \AA^2 (11.78 \AA^2) for two- and three-layer graphene, respectively. In Fig. 4 we show both the WFs and the band structure in two-layer graphene. The interpolated bands (red circles) perfectly match the DFT bands (solid line).

In the other applications demonstrated in the following part of this work, we are more interested in WFs as an efficient tight-binding basis rather than their actual shape and chemical interpretation. In such cases, we are interested in obtaining a set of well-localized WFs accurately describing single-particle states over a potentially wide energy range. This task leads to the question how to define appropriate energy windows. First, we fix the inner energy window to an interval in which we wish to obtain an accurate description

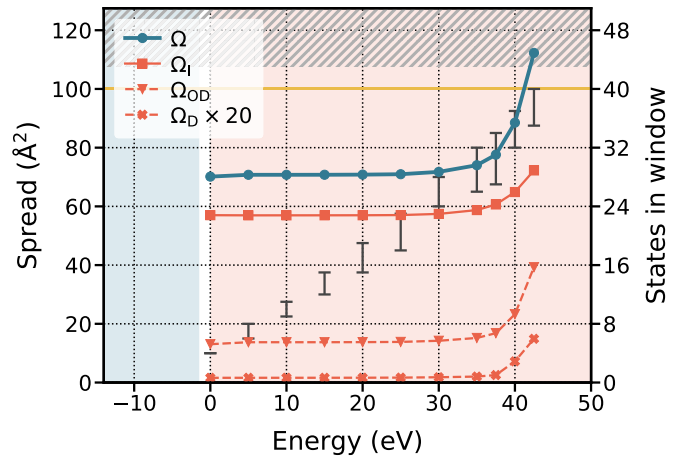


FIG. 5. Total spread Ω (blue circles) and its components (red squares, triangles, and crosses) of 40 WFs in GaP obtained for different choices of the inner energy window. The outer window is fixed to the displayed energy region covering between 43 and 51 bands (right vertical axis, gray hatched area) including the four valence bands (blue shaded area) and conduction bands (red shaded area) up to 50 eV above the CBM. Each data point corresponds to a different upper boundary of the inner window, whereas the lower boundary is kept fixed at the bottom of the valence bands. The gray bars display the number of bands (range over all \mathbf{k} points) that fall inside the inner window.

of electronic states in terms of localized WFs. This range may vary between different materials and also depends on the quantities of interest. Once the inner window is set, both the outer window and the number of WFs to be disentangled must be chosen. The latter must not exceed the minimal number of bands in the outer window and must not be smaller than the maximum number of bands in the inner window (minimum/maximum over all \mathbf{k} points).

To investigate the influence of the energy windows on the localization of the WFs, we disentangle 40 WFs from a manifold of both valence and conduction bands in GaP. We fix the outer energy window such that it covers between 43 and 51 bands including the four valence bands and conduction bands up to 50 eV above the CBM (see Fig. 5). The lower boundary of the inner window is fixed at the bottom of the valence bands and we vary the upper boundary. If the upper boundary is at 0 eV, the inner window contains only the four valence bands. The upper considered limit of 42.5 eV corresponds to the inner window containing between 35 and 40 bands. A further increase of the upper boundary is not meaningful because the inner window must not contain more bands than the number of requested WFs (set to 40 in this example). Figure 5 shows that the gauge-independent part of the spread Ω_1 which is the result of the disentanglement procedure smoothly decreases with a decreasing width of the inner window. This is a result of an increasing variational freedom in finding an optimal subspace and comes from the fact that less bands are forced to be part of that subspace. However, although the number of states that are fixed in the inner window decreases almost linearly with the width of the inner window, Ω_1 quickly converges toward a constant value and does not further decrease by releasing more bands from the inner window. The diagonal

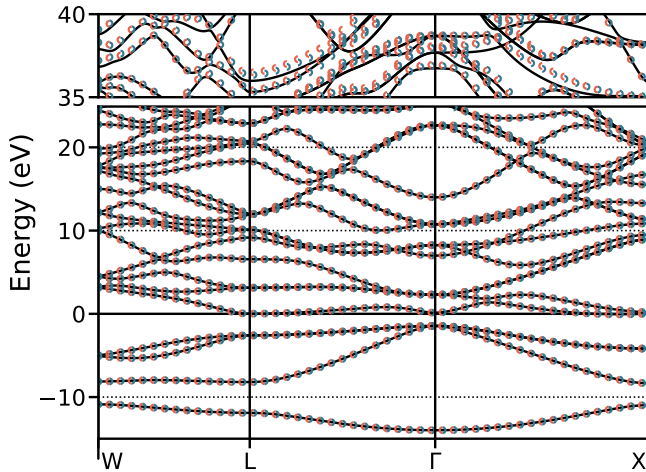


FIG. 6. Electron band structure in GaP from DFT (solid line) and Wannier interpolation for different inner energy windows. The outer window ranges from the bottom of the valence bands up to 50 eV above the CBM. The inner window spans the four valence bands (red right open circles) and additionally conduction bands up to 25 eV (blue left open circles), respectively.

and off-diagonal parts of the spread and hence the total spread show a similar behavior. From this observation, we conclude that well-localized WFs can be constructed over a wide energy range and that the energy interval in which they accurately describe electronic states can exceed the inner window (in which the Wannier bands are pinned down to the exact energies) by far. Indeed, we find no noticeable difference between Wannier band structures obtained with the upper inner-window boundary set to 0 and 25 eV. Moreover, both match the DFT band structure very well (see Fig. 6).

We derive the following criteria for the construction of well-localized WFs suitable for accurate interpolation purposes over wide energy ranges: (1) The number of WFs that are to be disentangled should be as close as possible to the number of full bands inside the outer window. (2) The difference between the inner and the outer windows should be as small as possible. (3) In practice, often the first two criteria must be loosened to some extent for the sake of convergence, i.e., to allow for enough variational freedom in the disentanglement procedure to eventually obtain a sufficiently smooth subspace. Otherwise, the following minimization of the spread may fail to converge. If we strictly obey (1) and (2), there will be many \mathbf{k} points at which all states must be part of the disentangled subspace leaving not much to disentangle and making the following minimization prone to failure. On the other hand, if we have a large difference between the inner and the outer windows and set the number of WFs close to the middle between both windows, we will have maximal freedom in the disentanglement, which will lead to a very strong reduction of Ω_I . However, the resulting WFs may be too localized in the sense that they will lead to unphysically flat bands inside the outer window which will also negatively influence the interpolation inside the inner window close to the window boundaries.

B. Interpolation of energy eigenvalues

The most obvious application of WFs is the interpolation of single-particle eigenenergies. For an arbitrary point \mathbf{q} in reciprocal space, the corresponding energies $\epsilon_n^{\mathbf{q}}$ are given as the eigenvalues of the Hamiltonian matrix from Eq. (5). In practice, \mathbf{q} is usually a point along a path connecting high-symmetry points in the BZ, when it comes to the calculation of band structures, or a point on a grid which is denser than the original grid on which the first-principles calculation was carried out. Such dense grids are often used to approximate integrations over the BZ by a discrete sum over a finite set of points. One key quantity of interest that involves such a BZ integration is the density of states (DOS). We use our implementation to investigate the joint DOS (JDOS) in aluminum. The JDOS is the phase-space contribution to optical excitations and can be calculated as

$$\text{JDOS}(\omega) = \int_{\text{BZ}} \sum_{o,u} \delta[\epsilon_u(\mathbf{k}) - \epsilon_o(\mathbf{k}) - \omega] d\mathbf{k}, \quad (11)$$

where o and u denote the occupied and unoccupied states for a given \mathbf{k} , respectively, and ω is the excitation energy. Note that the JDOS divided by ω^2 is proportional to the independent-particle optical spectrum with constant transition matrix elements. The spectrum of metals such as Al can be described well within the independent-particle picture since excitonic effects play a minor role due to the effective screening. Earlier calculations of optical spectra in Al showed that very dense integration grids containing several thousand irreducible \mathbf{k} points are needed to obtain convergence of the spectra [12,13]. In particular, also a strong dependence of the peak positions was observed [13]. To investigate the influence of the BZ grid on the JDOS in Al, we perform a DFT calculation within the generalized gradient approximation (GGA) using the PBE xc functional [14] on a $12 \times 12 \times 12$ \mathbf{k} mesh. From an outer (inner) energy window of -15 to 80 eV (-15 to 40 eV) 25 MLWFs are constructed using the disentanglement procedure. Hereby, the zero-energy point corresponds to the Fermi level. The OPFs are obtained from 34 trial orbitals. The structural and computational parameters used in this and all other calculations can be found in Tables III and IV, respectively. We interpolate the eigenvalues on different uniform integration grids by the use of MLWFs, and the improved tetrahedron method [15] is employed to evaluate the integral in Eq. (11). The resulting JDOS is shown in Fig. 7. We observe both a redshift and a significant sharpening of the two major peaks in the investigated energy region. Both peaks eventually converge at around 0.6 and 1.6 eV for 120^3 and 80^3 \mathbf{k} points, respectively. We notice that it is more difficult to achieve convergence in the low-energy region. For energies below 0.4 eV grids with more than 200^3 uniformly spaced points are needed (solid red line). The position of the two peaks around 0.6 and 1.6 eV in the converged curve are in excellent agreement with earlier calculations of the JDOS [12,16] as well as calculations [12,13,16] and measurements [16] of optical spectra. The redshift and sharpening with increasing grid densities was also found in calculations of optical spectra [13], where obviously the ratio of the peak heights differs from the optical spectra since transition probabilities are not taken into account in the JDOS.

TABLE III. Structural parameters for all materials investigated in Secs. III B–III E.

Compound	Lattice parameters						Lattice type	Space group	Atoms in unit cell
	a (Å)	b (Å)	c (Å)	α (deg)	β (deg)	γ (deg)			
Al	2.838	2.838	2.838	60	60	60	Cubic	$Fm-3m$	1Al
Ga ₂ O ₃	6.302	6.302	5.807	76.6	103.4	152.1	Monoclinic	$C2/m$	4Ga, 6O
TiO ₂	4.638	4.638	2.969	90	90	90	Tetragonal	$P4(2)/mmm$	2Ti, 4O
ZrS ₂	3.66	3.66	20.128	90	90	60	Hexagonal	$P3m1$	1Zr, 2S
Py@ZnO	6.310	6.310	23.284	90	90	112.3	Triclinic	$P1$	5C, 5H, 1N, 16Zn, 16O

C. Accuracy of energy interpolations

The claim that the WFs constructed according to Eq. (1) form an equivalent description of the subspace spanned by the Bloch states $\psi_{n,\mathbf{k}}$ under consideration only holds rigorously for isolated groups and in the limit of an exact BZ integral. In practice, however, the BZ is sampled by a finite set of points. As a result, the computed WFs become periodic with respect to a supercell conjugate to the BZ grid. This can lead to a nonvanishing overlap between a WF and its periodic images in neighboring supercells and ruins the exactness of the tight-binding basis from Eq. (4) which is given in the limit of an exact BZ integral. In turn, the interpolated eigenvalues at some point \mathbf{q} that does not belong to the original first-principles grid deviates from the exact solution. Note that *exact* is meant within the limitations of the first-principles calculations, i.e., the interpolated energy deviates from the result one would obtain by directly performing the calculations at the point \mathbf{q} . MLWFs associated with isolated bands are reported to be exponentially localized [17]. This claim was proved for insulators with time-reversal symmetry [18]. As a consequence, we expect the overlap with supercell images and thus the error in the interpolation to decay exponentially with increasing grid size. To investigate this behavior for the materials studied in this work, we proceed as follows. We consider a set of different grids $\{\mathbf{k}\}_1, \dots, \{\mathbf{k}\}_n$ (ordered with increasing grid density) for which we want to predict the accuracy of interpolated eigenenergies. First, we compute the self-consistent KS potential and electron density on the densest grid under consideration $\{\mathbf{k}\}_n$. This self-consistent density serves as a starting point for further calculations. We use it to obtain the eigenvalues $\hat{\epsilon}_n^{\mathbf{q}}$ on a much denser interpolation grid $\{\mathbf{q}\}$ by a non-self-consistent diagonalization of the KS Hamiltonian.

The dense interpolation grid is chosen to be shifted to ensure a sampling on inequivalent points. This set of energies $\hat{\epsilon}_n^{\mathbf{q}}$ forms the reference to which we compare the interpolated energies. Now, for each of the grids $\{\mathbf{k}\}_1, \dots, \{\mathbf{k}\}_{n-1}$ both wave functions and eigenenergies are calculated non-self-consistently starting from the density obtained on the grid $\{\mathbf{k}\}_n$. Lastly, for all grids $\{\mathbf{k}\}_1, \dots, \{\mathbf{k}\}_n$ MLWFs are constructed and used to interpolate the eigenvalues onto the dense shifted interpolation grid $\{\mathbf{q}\}$. The interpolated energies are denoted by $\epsilon_n^{\mathbf{q}}$. For each grid, we compute the interpolation error as the root-mean-square deviation of the interpolated energies from the calculated reference energies:

$$\delta\epsilon_{\text{RMS}} = \sqrt{\frac{1}{JN_{\mathbf{q}}} \sum_{n,\mathbf{q}} (\epsilon_n^{\mathbf{q}} - \hat{\epsilon}_n^{\mathbf{q}})^2}. \quad (12)$$

In order to compare BZ samplings for systems with different unit-cell size and dimensionality, we introduce the linear \mathbf{k} -point density which is given by $(N_{\mathbf{k}}/V_{\text{BZ},d})^{1/d}$, where $N_{\mathbf{k}}$ is the total number of nonreduced \mathbf{k} points, d is the dimensionality of the system, and $V_{\text{BZ},d}$ is the volume of the corresponding d -dimensional BZ.

We carry out DFT calculations for various materials using the PBE xc functional and follow the procedure described above. The results are presented in Fig. 8. The graphs indicate that an exponential decay of the interpolation error is an overall suitable assumption for most of the systems studied within this work. It is even found for the interpolation of entangled bands (empty circles, dashed lines) although there is no reason to assume an exponential localization of WFs obtained from the disentanglement procedure. The exponential decay is observed particularly well in the case of β -Ga₂O₂ for both the valence and the conduction bands. For TiO₂, however, the

TABLE IV. Computational parameters for the construction of OPFs and MLWFs in all materials investigated in Secs. III B–III E.

Compound	Occupation	Method	Wannier functions	Trial orbitals	Angular character of LOs	Unoccupied states per atom
Al	Occupied + unoccupied	Disentangle	25	34	s, p, d, f, g	20
TiO ₂	Occupied	Isolated	12	30	s, p, d	10
	Unoccupied	Isolated	10	26	s, p, d	10
	Unoccupied	Disentangle	148	454	s, p, d, f	10
Ga ₂ O ₃	Occupied	Isolated	18	60		10
	Unoccupied	Disentangle	37	90		10
ZrS ₂	Occupied	Isolated	6	21	s, p, d	10
	Unoccupied	Disentangle	3	18	s, p, d	10
	Unoccupied	Disentangle	27	34	s, p, d, f	10
Py@ZnO	Unoccupied	Disentangle	60	159	s, p, d	10

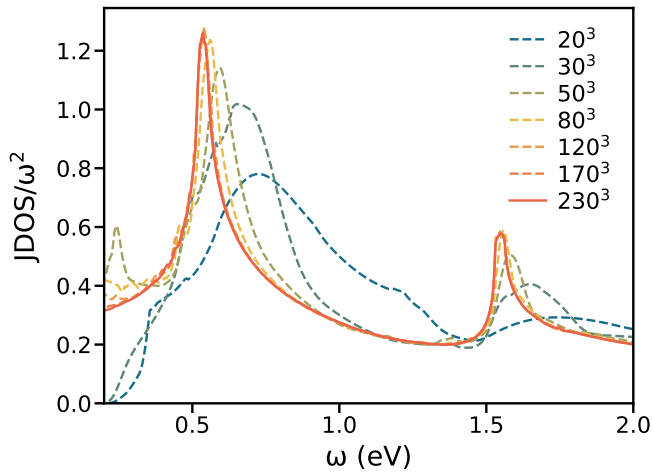


FIG. 7. Joint density of states for aluminum obtained from Wannier interpolation on different BZ integration grids ranging from $20 \times 20 \times 20$ to $230 \times 230 \times 230$.

behavior differs considerably from a pure exponential decay. Similar investigations have been performed before for a set of isolated bands in lead and for entangled bands in lithium [19]. There, the same behavior of a decreasing rate of decay for increasing grid densities (as it is clearly visible for TiO_2 in our calculations) was observed. Further, it was shown for one-dimensional (1D) systems that the localization of energy matrix elements follows a power law times an exponential [17]. Such a model also fits well to our results obtained for two- and three-dimensional systems. For all systems studied, an interpolation accuracy in the meV regime can be reached with manageable grid densities. Going to higher accuracies, however, will require higher grid densities than presented in Fig. 8 which may be feasible for KS-DFT eigenvalues but become rather cumbersome for the interpolation of generalized

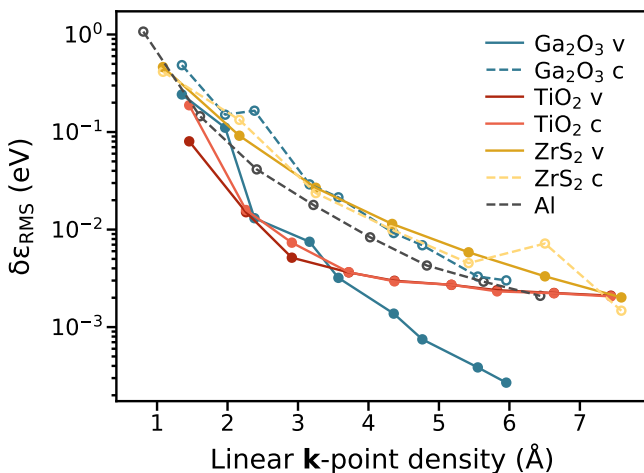


FIG. 8. Wannier interpolation error as a function of the \mathbf{k} -grid density. The filled (empty) circles mark the error of interpolated eigenvalues obtained from MLWFs representing isolated (entangled) bands in various systems for both valence (v) and conduction (c) bands. The lines serve as guides to the eye.

KS eigenvalues obtained from hybrid xc functionals or quasiparticle energies obtained from the GW approach.

D. Effective masses and band extrema

The accurate and inexpensive energy interpolation using WFs allows for a systematic search for band extrema. In semiconductors, the most interesting extremal points of the energy dispersion $\epsilon_n(\mathbf{k})$ typically are the highest occupied state (valence band maximum, VBM) and the lowest unoccupied state (conduction band minimum, CBM) determining the band gap and its type (direct or indirect). Finding their position is challenging when they are not located at a high-symmetry point in the BZ. In this case, they are usually not contained in the uniform BZ sampling employed in the DFT calculation. We use our implementation to determine the exact position of the VBM and CBM in $\beta\text{-Ga}_2\text{O}_3$ [20], focusing on the effect of different xc treatments and levels of theory. To this extent, the KS equations are solved within the LDA parametrized by Perdew and Wang [21], GGA using PBEsol [22], and the nonlocal hybrid functional PBE0 with 25% of Hartree-Fock exchange [23]. Furthermore, quasiparticle self-energy corrections to the PBEsol eigenvalues are computed using the G_0W_0 approximation. The (generalized) KS calculations are carried out using $8 \times 8 \times 4$ \mathbf{k} points in the full BZ. In the G_0W_0 calculation, a $4 \times 4 \times 4$ \mathbf{k} mesh and all empty states are used following the prescription in Ref. [10]. The set of 18 valence bands is transformed into MLWFs using the algorithm for isolated groups. The spread Ω of the initial guess obtained using the OPF method is only 1% larger than the final minimum for all xc treatments. The WFs describing the conduction bands are obtained by the disentanglement procedure using an outer (inner) energy window of 30 eV (20 eV) above the Fermi level which was set to the middle of the band gap.

The CBM in $\beta\text{-Ga}_2\text{O}_3$ is known to be located at the zone center Γ . This is confirmed by our calculations. The position of the VBM, however, is not at one of the high-symmetry points in the BZ. It was reported to be on the line connecting the high-symmetry points L and I [24] [see Fig. 9(a)] which is in accordance with our findings. In Fig. 9(b), we show the highest valence band and the position of the VBM for the different theoretical approaches. We find a weak dependence of the exact position of the VBM on the used xc treatment. Although the position is nearly the same for LDA and PBEsol, it is slightly closer to L for the hybrid functional and the G_0W_0 calculation. The exact values are reported in Table V by the parameter β varying from 0 to 1 between the points L and I.

Further in Table V, we present the resulting band gaps. A comparison of the Γ - Γ gap with the experimental gap of about 4.9 eV [25,26] reveals that the nonlocal hybrid functional yields the best agreement with a direct gap of 5.0 eV. As expected, the (semi)local functionals LDA and PBEsol severely underestimate the gap. Also, the quasiparticle gap of 4.5 eV is underestimated. However, none of the theoretical values consider band renormalization effects due to electron-phonon interaction which makes a direct comparison with experimental results difficult. In all cases, the indirect gap is about 30 meV smaller than the Γ - Γ gap.

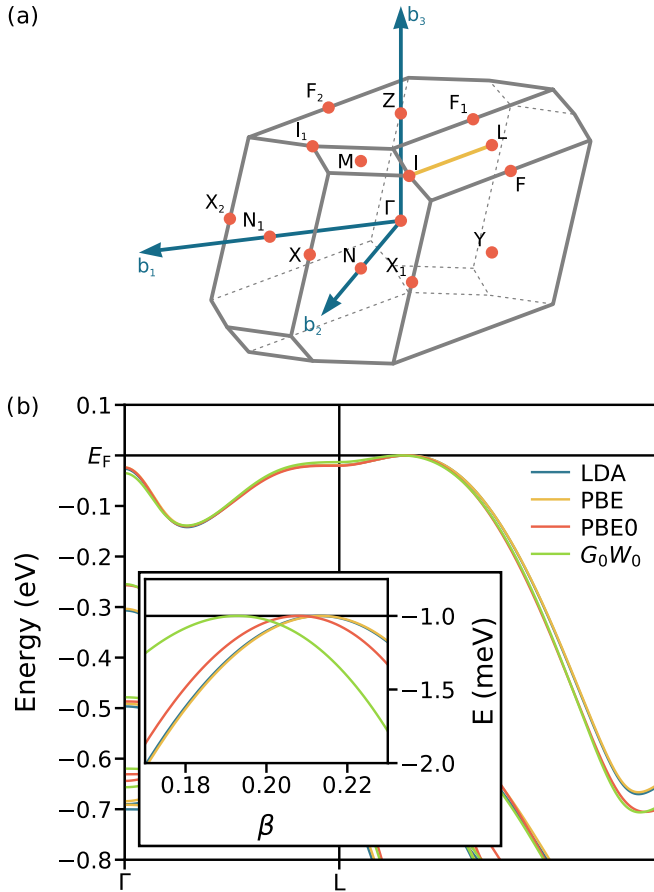


FIG. 9. (a) Brillouin zone of β -Ga₂O₃. The line on which the valence band maximum (VBM) is found is highlighted in yellow. (b) Highest valence band in β -Ga₂O₃ along the high-symmetry lines Γ -L and L-I for different theoretical approaches. The Fermi level is set to the VBM. Inset: region around the VBM. The parameter β describes its position between the high-symmetry points L ($\beta = 0$) and I ($\beta = 1$).

TABLE V. Position of the band extrema, effective masses (in units of m_0), and fundamental band gaps (in eV) in β -Ga₂O₃ for different xc treatments determined analytically using Wannier interpolation. The parameter β describes the position of the VBM along the line between the high-symmetry points L ($\beta = 0$) and I ($\beta = 1$).

	LDA	PBE	PBE0	$G_0W_0@PBE$
VBM				
β	0.2132	0.2136	0.2081	0.1953
m_{xx}^*	2.94	2.95	2.97	3.20
m_{yy}^*	3.15	3.14	2.90	3.41
m_{zz}^*	4.30	4.39	4.73	3.02
m_{xz}^*	0.232	0.258	0.572	0.089
CBM				
m_{xx}^*	0.238	0.234	0.275	0.294
m_{yy}^*	0.263	0.263	0.280	0.333
m_{zz}^*	0.253	0.251	0.273	0.280
E_g (eV)				
Indirect	2.271	2.290	5.009	4.490
Γ - Γ	2.297	2.314	5.033	4.525

The simple form of the single-particle wave functions expressed in the WF tight-binding basis [see Eq. (4)] allows for an analytic expression of \mathbf{k} -space derivatives since the dependence on the wave vector only comes from the Fourier phase factor while the set of WFs is the same for all \mathbf{k} points. Thus, the use of finite differences or numerical fitting methods (which are usually used to calculate derivatives) can be avoided. This analytical approach allows for the direct calculation of the particle group velocity

$$\mathbf{v}_n(\mathbf{k}) = \nabla_{\mathbf{k}} \epsilon_n(\mathbf{k}) \quad (13)$$

and the effective-mass tensor

$$\mathbf{m}_n^*(\mathbf{k}) = [\nabla_{\mathbf{k}} \nabla_{\mathbf{k}}^T \epsilon_n(\mathbf{k})]^{-1}. \quad (14)$$

Note that atomic units are used in Eqs. (13) and (14), and $\nabla_{\mathbf{k}}$ is a column vector. We follow the derivations by Yates and co-workers [19] in order to evaluate the analytic expression of the first and second \mathbf{k} derivative of the band dispersion in β -Ga₂O₃ to determine the effective masses at the CBM and VBM. The results are given in Table V. We find $\mathbf{m}_{\text{CBM}}^*$ to be almost diagonal and isotropic. The electron effective mass varies from 0.237 to 0.333 electron rest masses depending on the direction and the xc treatment. Again, there are no noticeable differences between LDA and PBEsol. For the hybrid functional PBE0 the CBM is more isotropic compared to LDA and PBEsol, and the electrons are slightly heavier with effective masses between 0.273 m_0 and 0.280 m_0 . These values are in perfect agreement with the (almost isotropic) value of 0.281 m_0 previously reported for the hybrid functional HSE06 [27]. The results for LDA and PBEsol are in good agreement with values of around 0.23 m_0 that were previously obtained for LDA [28]. For the quasiparticles we find even higher effective masses between 0.280 m_0 and 0.333 m_0 .

At the VBM, the effective-mass tensor takes the following form:

$$\mathbf{m}_{\text{VBM}}^* = \begin{pmatrix} m_{xx}^* & 0 & m_{xz}^* \\ 0 & m_{yy}^* & 0 \\ m_{xz}^* & 0 & m_{zz}^* \end{pmatrix}, \quad (15)$$

where the m_{xy}^* and m_{yz}^* components do not vanish completely but are about three orders of magnitude smaller than the diagonal components and therefore neglected. According to our calculations, the VBM is more anisotropic. For LDA, PBEsol, and PBE0, we obtain similar hole effective masses in the x and y directions of around 3 m_0 and values between 4.3 m_0 (LDA) and 4.7 m_0 (PBE0) in the z direction. The quasiparticle calculation differs noticeably from the other three approaches and suggests heavier holes in the x and y directions and lighter holes in the z direction. Overall, our results are comparable with those of Ref. [28] for the y and z directions but differ noticeably in the x direction for which Ref. [28] reported a hole effective mass of $m_{xx}^* = 6.14 m_0$ which is about twice the value we find. However, both the exact position of the VBM and the band curvature are difficult to determine accurately due to the very low dispersion in the valence band top region and the occurrence of multiple maxima that differ only little in energy. For instance, there is another maximum at Γ only 30 meV below the VBM (see Fig. 9). We are not aware of any reports on experimental hole effective masses in β -Ga₂O₃ to compare with.

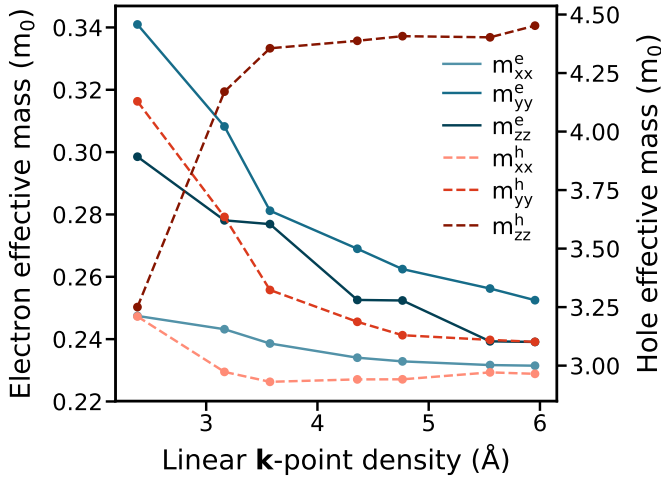


FIG. 10. Diagonal effective masses for electrons (blue, solid lines) and holes (red, dashed lines) in β -Ga₂O₃ obtained from Wannier interpolation on top of PBE as a function of the first-principles \mathbf{k} -grid density.

In order to estimate the accuracy of the determined effective masses, we perform a similar convergence test as it is done in Sec. III C for the predicted energies. In Fig. 10 we present the diagonal entries of the effective-mass tensor for electrons at the CBM (blue, solid lines) and holes at the VBM (red, dashed lines) obtained from the analytic approach starting from PBEsol calculations on different \mathbf{k} grids. The values presented in Table V (with the exception of G_0W_0) are obtained on a grid corresponding to a linear \mathbf{k} -point density of about 4.8 Å⁻¹. Figure 10 shows that for this grid density the hole effective masses are almost converged and we estimate an uncertainty of about 0.1 m_0 ($\approx 3\%$). In contrast, the noticeably smaller electron effective masses are much harder to predict accurately. They are not yet fully converged in the studied range of \mathbf{k} -point densities and thus we estimate a larger relative uncertainty for the numbers in Table V of about 0.02 m_0 ($\approx 10\%$).

E. Interpolation of wave functions

The diagonalization of the Wannier-interpolated Hamiltonian $\mathcal{H}_{mn}^{\mathbf{q}}$ gives also rise to the interpolated wave functions. They are expressed in the form

$$\psi_{n,\mathbf{q}}(\mathbf{r}) = \sum_m V_{mn}^{\mathbf{q}} \phi_{m,\mathbf{q}}(\mathbf{r}), \quad (16)$$

where $V_{mn}^{\mathbf{q}}$ is the eigenvector of $\mathcal{H}^{\mathbf{q}}$ corresponding to the eigenvalue $\epsilon_n^{\mathbf{q}}$, and $\phi_{m,\mathbf{q}}$ is defined by Eq. (4). The analysis of these wave functions offers deeper physical and chemical insights. To this extent, we decompose $\psi_{n,\mathbf{q}}$ in particular atomic states by an expansion in a series of spherical harmonics Y_{lm} times radial functions $\varphi_{n,\mathbf{q},lm}^{\alpha}(r)$ inside the individual muffin-tin spheres α :

$$\psi_{n,\mathbf{q}}^{\alpha}(\mathbf{r}) = \sum_l \sum_{m=-l}^l \varphi_{n,\mathbf{q},lm}^{\alpha}(|\mathbf{r} - \mathbf{R}_{\alpha}|) Y_{lm}(\widehat{\mathbf{r} - \mathbf{R}_{\alpha}}). \quad (17)$$

Within the (L)APW + LO basis, this expansion is straightforward. Now, we calculate the contribution of the state

$\psi_{n,\mathbf{q}}$ to the number of electrons inside the muffin-tin sphere α with radius R_{α} by integrating the partial density $\rho_{n,\mathbf{q}}(\mathbf{r}) = |\psi_{n,\mathbf{q}}(\mathbf{r})|^2$:

$$\begin{aligned} \int_{\text{MT}_{\alpha}} \rho_{n,\mathbf{q}}(\mathbf{r}) d\mathbf{r} &= \sum_l \sum_{m=-l}^l \int_0^{R_{\alpha}} r^2 |\varphi_{n,\mathbf{q},lm}^{\alpha}(r)|^2 dr \\ &= \sum_l b_{n,\mathbf{q}}^{\alpha,l}. \end{aligned} \quad (18)$$

The second line of Eq. (18) defines the band character $b_{n,\mathbf{q}}^{\alpha,l}$ which is interpreted as the contribution of electrons with angular character l and wave vector \mathbf{q} inside the muffin-tin sphere α to the n th energy band.

We interpolate the band character for TiO₂ in the rutile structure and for a monolayer of the two-dimensional (2D) material ZrS₂. The calculation of TiO₂ is carried out using the hybrid xc functional PBE0 and a $6 \times 6 \times 9$ \mathbf{k} -point grid. The 12 valence bands and the 10 lowest conduction bands are transformed into MLWFs separately using the algorithm for isolated bands. Again, the spread Ω of the initial guess is only 1% and 2% off the final minimum for the two groups, respectively. For the higher conduction bands, 148 WFs are disentangled using an outer (inner) energy window of 8 to 130 eV (8 to 76 eV), which demonstrates the ability to find localized WFs accurately describing electron states up to very high energies. In the case of ZrS₂, quasiparticle energies are calculated within the G_0W_0 approximation on top of PBE for $8 \times 8 \times 1$ \mathbf{k} points. The six valence bands are treated as an isolated group. We disentangle the three Zr d -like bands which intersect with higher-energy conduction bands around the Γ point from the energy window between 0 and 4.75 eV. The initial guess for both the valence bands and the three disentangled conduction bands is only 2% larger than the final minimum. The remaining conduction bands are represented by 27 WFs disentangled from an outer (inner) energy window of 3.75 to 20 eV (4.75 to 10 eV). In the top panels of Figs. 11 and 12, we present the interpolated band structure and DOS for TiO₂ and ZrS₂, respectively. For obtaining the DOS, the energies and the band characters are interpolated on a grid of $60 \times 60 \times 90$ and $300 \times 300 \times 1$ points in the BZ for TiO₂ and ZrS₂, respectively. The bands and the DOS are colored according to the band character, i.e., the contribution of electrons from different atom species and with different angular character. Since the band character does not account for contributions from the interstitial region, the sum of the projected DOS (colored area) differs from the total DOS (black solid line). In the case of TiO₂, the 12 valence bands almost entirely originate from oxygen p -like states. The 12 symmetry-equivalent WFs corresponding to this group of bands [one illustrated in Fig. 11(b)] clearly reflect this character. The same holds for the isolated group of the 10 lowest conduction bands which exhibit dominantly titanium d character with some admixture of oxygen p -like states. Again, this is clearly reflected in the corresponding Wannier functions [Fig. 11(c)]. A similar behavior can be found in ZrS₂. The valence bands show a strong sulfur p character since the corresponding Wannier functions [Fig. 12(b)] are almost purely p like and centered at sulfur atoms. The Wannier functions corresponding to the three zirconium d -like bands

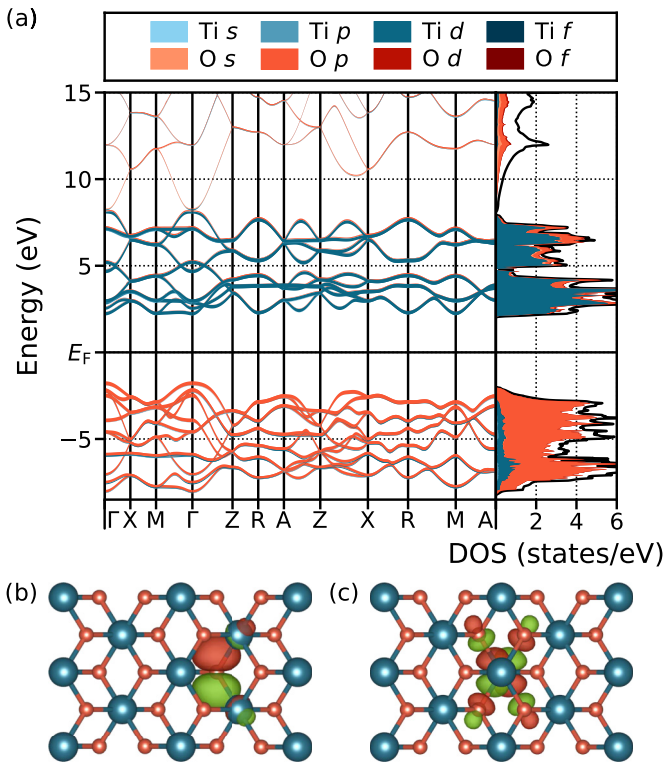


FIG. 11. Wannier interpolated band structure and DOS (a) for TiO_2 in the rutile structure calculated using PBE0. The Fermi level is set to the middle of the gap. The different shades of colors display the individual contributions of the wave functions at titanium (blue) and oxygen (red) atoms with different angular character (l). MLWFs corresponding the valence (b) and lowest conduction (c) bands. Note that the MLWFs are real valued. Positive (negative) isosurfaces are displayed in red (green).

in the lower conduction band region [Fig. 12(c)] clearly reflect the dominant Zr d character but also show contributions from sulfur p -like states.

Within the G_0W_0 approximation, a self-energy correction to the KS eigenvalues is calculated in order to obtain the quasiparticle energies. Often, these corrections (obtained on a uniform \mathbf{k} grid) are used to deduce a rigid scissors shift from which the band structure is then obtained. This approach, however, is not always justified, like for instance in hybrid inorganic-organic systems. The prototypical compound shown here consists of pyridine molecules chemisorbed on the $(10\bar{1}0)$ surface of a ZnO slab with 43 atoms in the unit cell (see bottom panel in Fig. 13). Using LDA KS states and eigenenergies as the starting point, the quasiparticle energies are computed non-self-consistently on $4 \times 4 \times 1$ \mathbf{k} points corresponding to a linear \mathbf{k} -point density of 3.9 \AA^{-1} . See Ref. [29] for further computational details. From an outer (inner) window of 13.6 eV (8.2 eV) above the Fermi level 60 WFs are disentangled to compute the quasiparticle band structure and compare it to the KS band structure (Fig. 13). Note that within the G_0W_0 approach the single-particle wave functions remain unchanged and just a quasiparticle correction to the eigenenergies is obtained. Hence, the MLWFs are also constructed from single-particle states. The quasiparticle energies are used for the interpolation according to Eqs. (4)

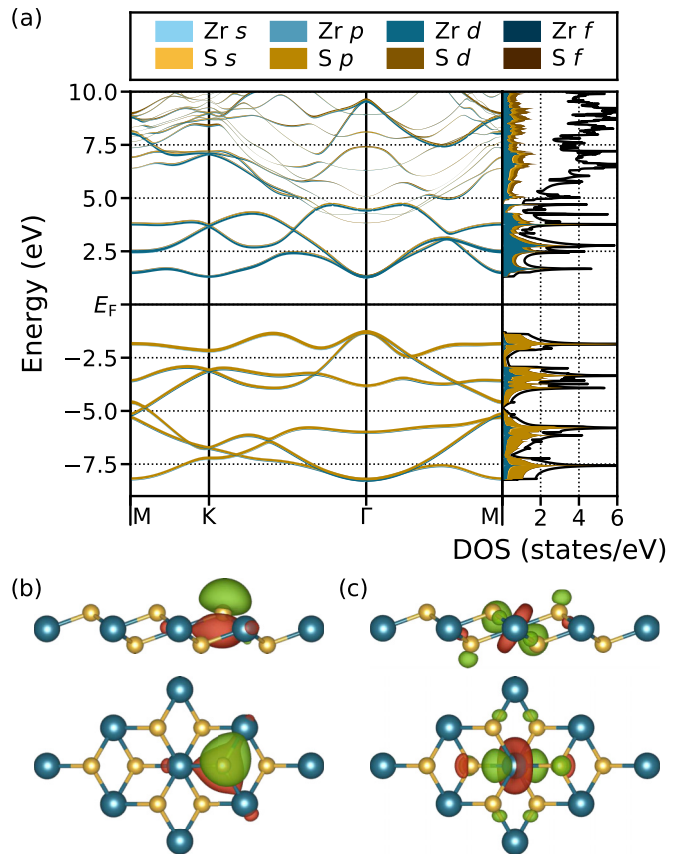


FIG. 12. Same as Fig. 11 for a ZrS_2 monolayer calculated using the G_0W_0 approximation on top of PBE. The upper (lower) illustration of the MLWFs shows the side (top) view.

and (5). The so-obtained wave functions are not to be confused with quasiparticle wave functions that one would obtain using the GW formalism with some flavor of self-consistency. Using the band character, we can attribute the individual energy bands to the constituents of the system. Bands displayed in blue are attributed to the organic molecule while red bands originate from the inorganic ZnO slab. Hybridized bands are colored in shades of green, yellow, and orange. In the bottom of Fig. 13, KS orbitals at Γ are shown, attributed to ZnO (red), pyridine (blue), and a hybridized state (yellow), respectively. It is evident that the quasiparticle self-energy correction has significantly different effects on the individual energy bands depending on their origin. While all conduction bands experience a general shift toward higher energies, the two flat molecular bands (blue) are subject to a much stronger upward shift with respect to the four parabolic ZnO bands (red). In contrast, the strongly hybridized band (yellow) is slightly shifted downward with respect to ZnO bands.

IV. CONCLUSIONS

We have presented an implementation of MLWFs within the (L)APW + LO method. By combining the well-established algorithm developed by Souza, Marzari, and Vanderbilt [2] with the more recently presented OPF technique [4], we are able to robustly construct MLWFs for various classes of materials without the need of projection functions

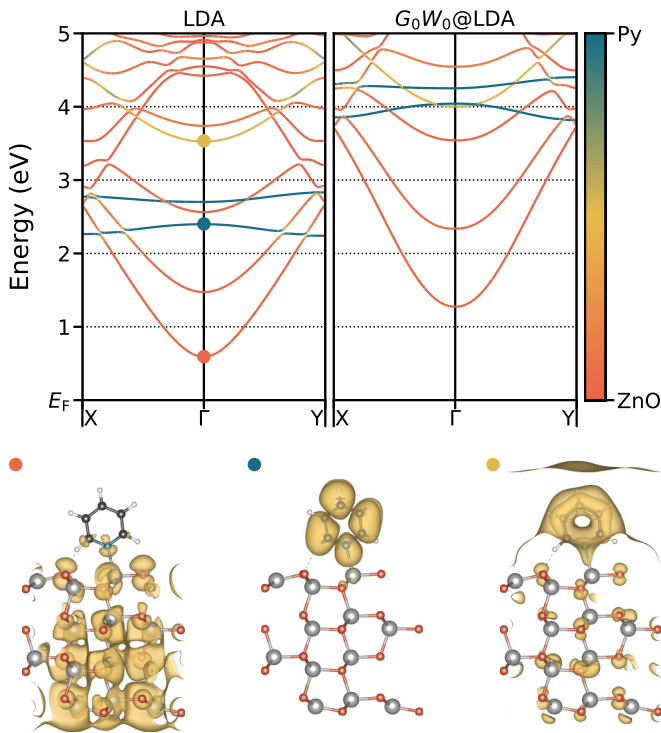


FIG. 13. Energy dispersion for the lowest KS and quasiparticle conduction bands in a hybrid inorganic-organic system (top left and top right, respectively). The bands are colored according to their origin. Bands attributed to the inorganic ZnO slab and the organic pyridine molecule are drawn in red and blue, respectively. The KS wave function for a hybridized state (yellow dot) as well as for states originating from ZnO (red dot) and pyridine (blue dot) are illustrated in the lower part.

being selected by the user. We use LOs as projection functions within the (L)APW + LO method. It is appealing due to its simplicity although they are strictly atom centered and vanishing in the interstitial region. This lack of flexibility can be overcome reliably by the use of the OPF approach. By automatically and systematically adding and selecting LOs from the pool of projection functions, we are able to calculate

MLWFs for both isolated and entangled bands in 2D and bulk semiconductors with small and medium sized unit cells, in (semi)metals as well as in complex hybrid systems containing an inorganic semiconductor and organic molecules.

This procedure gives access to accurate band structures and DOS based on more sophisticated methods such as generalized hybrid KS-DFT or quasiparticle calculations which otherwise would not be available due to the immense computational cost these methods come with. The same holds for other quantities that can be derived from the band structure directly such as band gaps, group velocities, and effective masses. According to our findings, a linear density of about four \mathbf{k} points per \AA^{-1} in reciprocal space in the underlying calculation suffices to predict electronic energies at an arbitrary point with an accuracy in the meV range. A deeper analysis of the interpolated wave function gives access to the band character and allows for a detailed interpretation of band structures and DOS. The results are in excellent agreement with calculations carried out in the original basis indicating that not just eigenenergies but also wave functions can be predicted accurately.

Future applications may involve MLWFs as basis functions in excited-state calculations using MBPT which often come with high effort simultaneously requiring dense \mathbf{k} grids. A reduction of the basis size and the simple access to wave functions and energies at arbitrary points in reciprocal space may help to reduce the computational cost of these approaches retaining the high precision of the (L)APW + LO method.

ACKNOWLEDGMENTS

This work was partially performed in the framework of GraFOx, a Leibniz-ScienceCampus supported by the Leibniz association. Parts of this work were funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Project No. 182087777-SFB 951. All input and output files can be downloaded from [30]. The LDA and G_0W_0 calculations of the Py@ZnO interface underlying our investigations were performed by O. Turkina, the G_0W_0 calculation of β -Ga₂O₃ by D. Nabok. We thank them for providing the data.

- [1] N. Marzari and D. Vanderbilt, *Phys. Rev. B* **56**, 12847 (1997).
- [2] I. Souza, N. Marzari, and D. Vanderbilt, *Phys. Rev. B* **65**, 035109 (2001).
- [3] A. Damle, L. Lin, and L. Ying, *J. Chem. Theory Comput.* **11**, 1463 (2015).
- [4] J. I. Mustafa, S. Coh, M. L. Cohen, and S. G. Louie, *Phys. Rev. B* **92**, 165134 (2015).
- [5] N. Marzari, A. A. Mostofi, J. R. Yates, I. Souza, and D. Vanderbilt, *Rev. Mod. Phys.* **84**, 1419 (2012).
- [6] G. H. Wannier, *Phys. Rev.* **52**, 191 (1937).
- [7] J. C. Slater and G. F. Koster, *Phys. Rev.* **94**, 1498 (1954).
- [8] A. Gulans, S. Kontur, C. Meisenbichler, D. Nabok, P. Pavone, S. Rigamonti, S. Sagmeister, U. Werner, and C. Draxl, *J. Phys.: Condens. Matter* **26**, 363202 (2014).
- [9] A. Gulans, A. Kozhevnikov, and C. Draxl, *Phys. Rev. B* **97**, 161105(R) (2018).
- [10] D. Nabok, A. Gulans, and C. Draxl, *Phys. Rev. B* **94**, 035118 (2016).
- [11] H. Abu-Farsakh and A. Qteish, *Phys. Rev. B* **75**, 085201 (2007).
- [12] K.-H. Lee and K. J. Chang, *Phys. Rev. B* **49**, 2362 (1994).
- [13] C. Ambrosch-Draxl and J. O. Sofo, *Comput. Phys. Commun.* **175**, 1 (2006).
- [14] J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [15] M. Kawamura, Y. Gohda, and S. Tsuneyuki, *Phys. Rev. B* **89**, 094515 (2014).
- [16] F. Szmulowicz and B. Segall, *Phys. Rev. B* **24**, 892 (1981).
- [17] L. He and D. Vanderbilt, *Phys. Rev. Lett.* **86**, 5341 (2001).

- [18] C. Brouder, G. Panati, M. Calandra, C. Mourougane, and N. Marzari, *Phys. Rev. Lett.* **98**, 046402 (2007).
- [19] J. R. Yates, X. Wang, D. Vanderbilt, and I. Souza, *Phys. Rev. B* **75**, 195121 (2007).
- [20] J. Furthmüller and F. Bechstedt, *Phys. Rev. B* **93**, 115204 (2016).
- [21] J. P. Perdew and Y. Wang, *Phys. Rev. B* **45**, 13244 (1992).
- [22] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke, *Phys. Rev. Lett.* **100**, 136406 (2008).
- [23] M. Ernzerhof and G. E. Scuseria, *J. Chem. Phys.* **110**, 5029 (1999).
- [24] H. Peelaers and C. G. Van de Walle, *Phys. Status Solidi B* **252**, 828 (2015).
- [25] M. Orita, H. Ohta, M. Hirano, and H. Hosono, *Appl. Phys. Lett.* **77**, 4166 (2000).
- [26] C. Janowitz, V. Scherer, M. Mohamed, A. Krapf, H. Dwell, R. Manzke, Z. Galazka, R. Uecker, K. Irmscher, R. Fornari, M. Michling, D. Schmeißer, J. R. Weber, J. B. Varley, and C. G. V. de Walle, *New J. Phys.* **13**, 085014 (2011).
- [27] J. B. Varley, J. R. Weber, A. Janotti, and C. G. Van de Walle, *Appl. Phys. Lett.* **97**, 142106 (2010).
- [28] K. Yamaguchi, *Solid State Commun.* **131**, 739 (2004).
- [29] O. Turkina, D. Nabok, A. Gulans, C. Cocchi, and C. Draxl, *Adv. Theory Simul.* **2**, 1800108 (2019).
- [30] See the NOMAD Repository, doi: 10.17172/NOMAD/2019.08.28-1.