Maximally-localized Wannier functions for entangled energy bands

Ivo Souza[1], Nicola Marzari[2], and David Vanderbilt[1]
[1] *Department of Physics and Astronomy, Rutgers University, Piscataway, New Jersey 08854-8019*
[2] *Department of Chemistry, Princeton University, Princeton, New Jersey 08544-1009*
(August 4, 2001)

We present a method for obtaining well-localized Wannier-like functions (WFs) for energy bands that are attached to or mixed with other bands. The present scheme removes the limitation of the usual maximally-localized WFs method (N. Marzari and D. Vanderbilt, Phys. Rev. B **56**, 12847 (1997)) that the bands of interest should form an isolated group, separated by gaps from higher and lower bands everywhere in the Brillouin zone. An energy window encompassing N bands of interest is specified by the user, and the algorithm then proceeds to disentangle these from the remaining bands inside the window by filtering out an optimally connected N-dimensional subspace. This is achieved by minimizing a functional that measures the subspace dispersion across the Brillouin zone. The maximally-localized WFs for the optimal subspace are then obtained via the algorithm of Marzari and Vanderbilt. The method, which functions as a postprocessing step using the output of conventional electronic-structure codes, is applied to the s and d bands of copper, and to the valence and low-lying conduction bands of silicon. For the low-lying nearly-free-electron bands of copper we find WFs which are centered at the tetrahedral interstitial sites, suggesting an alternative tight-binding parametrization.

PACS: 71.15.-m, 71.15.Ap, 71.20.-b

I. INTRODUCTION

When studying electrons in solids, it is often the case that only a small subset of the available one-electron states contributes significantly to the properties under consideration. Moreover, the states of interest typically lie within a limited energy range. For instance, for modeling electron transport or magnetic properties, only the partially filled bands close to the Fermi energy E_F are needed. This is the rationale behind the tight-binding and Hubbard models, in which only a few energy bands are kept.[1,2] Those models rely on the existence of a minimal set of spatially localized orbitals spanning the manifold of relevant states.

In recent years there has been growing interest in explicitly constructing such orbitals from first-principles density-functional calculations. One potential application consists in obtaining the parameters in correlated Hamiltonians by constraining the occupation of the orbitals to find the energy cost of deviating from the mean-field solution ("constrained density-functional theory"[3,4]). Another arises in the context of the "dynamical mean-field theory" which, when combined with density-functional methods, requires the specification of localized orbitals describing the narrow bands of interest.[5]

Wannier functions[6] (WFs) are a very natural type of localized orbital for extended systems. They play a central role in formal discussions of the tight-binding[1] and Hubbard[2] models. Traditionally they have often been invoked – although rarely calculated explicitly – as a convenient basis for describing local phenomena, such as impurities,[7] excitons,[7] and magnetic properties.[8] More recently, WFs have found important applications in connection with linear-scaling algorithms for electronic structure calculations.[9] Moreover, they play an important role in the theory of electronic polarization and localization in insulators, with the former quantity being related to the centers of charge of the WFs[10,11] and the latter to their quadratic spreads.[12,13] These developments have also led to generalizations of the concept of Wannier functions to correlated electron systems.[13–15]

The main obstacles to the construction of WFs in practical calculations have been their nonuniqueness (or "gauge dependence") and the difficulties in dealing with degeneracies among the Bloch states. These have been overcome by the development by Marzari and Vanderbilt of a general and practical method for extracting "maximally-localized" WFs from an isolated group of bands.[16] (By "isolated" we mean a group of bands that may become entangled with one another across the Brillouin zone, but is separated from all other bands by finite gaps throughout the entire Brillouin zone. The set of valence bands of an insulator constitutes an important example.) The method has been successfully used to describe the dielectric properties of several insulating systems, such as crystalline[16] and amorphous[17] semiconductors, ferroelectric perovskites,[18] liquid water,[19] compressed solid hydrogen,[20] and manganese oxide.[21] It has been implemented for plane-wave,[16] linear augmented plane-wave,[21] and tight-binding[20] basis sets.

However, in many cases the group of bands of interest is not isolated in the above sense, especially when dealing with metals or with the empty bands of insulators. For example, the conduction s band of an alkali metal is attached at points or lines of high symmetry to higher bands; the d bands of a noble or transition metal are hybridized with an s band, which in turn is attached to

higher bands; the conduction bands of a copper-oxide superconductor emerge from a dense group of bands below; and the four low-lying antibonding bands of a tetrahedral semiconductor are connected to higher conduction bands.

A successful technique that has been applied for constructing localized orbitals that describe such bands is the "downfolding" technique[22,23] that has been developed for electronic structure methods based on muffin-tin orbitals. There have also been previous attempts at constructing WFs for non-isolated groups of bands, namely for noble and transition metals[24–27] and for tetrahedral semiconductors.[28,29] These attempts fall into two categories: (i) the WFs are obtained directly from a variational principle, as suggested by Kohn,[30] or (ii) they are obtained as Fourier transforms of Bloch functions, with the help of a model Hamiltonian that reproduces the band structure in the desired energy range, as suggested by Bross.[31]

We will describe an alternative Wannier-based approach that is closer in spirit to the Fourier transform method of Bross and co-workers, but does not require the construction of an auxiliary model Hamiltonian. The method can be regarded as an extension to the case of attached bands of the maximally-localized WF method of Marzari and Vanderbilt.[16] It has the desirable features that it can be implemented with any basis set (e.g., plane waves), and requires minimal user-intervention (the only "adjustable parameter" being a specification of the energy range of interest). Like the approach of Ref. 16, ours is a "postprocessing" method, taking as its input the Bloch eigenstates and eigenvalues calculated by a standard electronic-structure code.

Strictly speaking, the resulting orbitals are not WFs (or even "generalized WFs"[16]) in the usual sense. They are nevertheless Wannier-like in the fundamental sense that they are obtained via an integral over the Brillouin zone of Bloch-like functions. As such they form an orthonormal, localized basis of the same Bloch subspace from which they were constructed.

The power of the present approach is illustrated by one particularly striking result that emerged from the work. In Sec. IV B 3 we find that a rather natural representation of the low-lying bands of an fcc metal like copper can be made in terms of a set of five Cu d-like WFs and two additional WFs centered at the tetrahedral interstitial locations. This provides a basis for a concise tight-binding representation of copper that has not, to our knowledge, previously been considered.

The paper is organized as follows. In Sec. II we review the method of Marzari and Vanderbilt for obtaining well-localized WFs for an isolated group of bands. In Sec. III we describe our procedure for dealing with attached energy bands, and in Sec. IV we illustrate it with a set of applications. Finally, in Sec. V we present a summary and conclusions.

II. MAXIMALLY-LOCALIZED WANNIER FUNCTIONS FOR AN ISOLATED GROUP OF BANDS

A set of WFs $w_{n\mathbf{R}}(\mathbf{r}) = w_n(\mathbf{r}-\mathbf{R})$ labeled by Bravais lattice vectors $\mathbf{R}$ can be constructed from the Bloch eigenstates $\psi_{n\mathbf{k}}$ of band n using the unitary transformation

$$w_{n\mathbf{R}}(\mathbf{r}) = \frac{v}{8\pi^3}\int_{\mathrm{BZ}} e^{-i\mathbf{k}\cdot\mathbf{R}}\,\psi_{n\mathbf{k}}\,d\mathbf{k}, \tag{1}$$

where v is the volume of the unit cell of the crystal and the integral is over the Brillouin zone. Except for the constraint $\psi_{n,\mathbf{k}+\mathbf{G}} = \psi_{n\mathbf{k}}$ for all reciprocal lattice vectors $\mathbf{G}$, the overall phases of the Bloch functions $\psi_{n\mathbf{k}} = e^{i\mathbf{k}\cdot\mathbf{r}}u_{n\mathbf{k}}$ are at our disposal. However, a different choice of phases (or "gauge"),

$$u_{n\mathbf{k}} \to e^{i\varphi_n(\mathbf{k})}\,u_{n\mathbf{k}}, \tag{2}$$

does not translate into a simple change of the overall phases of the WFs; their shape and spatial extent will in general be affected. If the band is isolated, Eq. (2) is the only allowed type of gauge transformation for changing the set of WFs $w_n(\mathbf{r}-\mathbf{R})$ associated with that band. In the case of an isolated group of N bands, the allowed transformations are of the more general form

$$u_{n\mathbf{k}} \to \sum_{m=1}^{N} U^{(\mathbf{k})}_{mn}\,u_{m\mathbf{k}}, \tag{3}$$

where $U^{(\mathbf{k})}$ is a unitary matrix that mixes the bands at wave vector $\mathbf{k}$. The resulting orbitals are called "generalized Wannier functions".[16]

Once a measure of localization has been chosen and an isolated group of bands specified, the search for the corresponding set of "maximally-localized" WFs becomes a problem of functional minimization in the space of the matrices $U^{(\mathbf{k})}$. The strategy of Ref. 16 consists in minimizing the sum of the quadratic spreads of the Wannier probability distributions $|w_n(\mathbf{r})|^2$,

$$\Omega = \sum_{n=1}^{N}\left(\langle r^2\rangle_n - \langle\mathbf{r}\rangle_n^2\right), \tag{4}$$

where the sum is over the chosen group of bands and $\langle\mathbf{r}\rangle_n = \int \mathbf{r}\,|w_n(\mathbf{r})|^2\,d\mathbf{r}$, etc. Interestingly, the resulting "maximally-localized" (or "maxloc") WFs turn out to be real, apart from an arbitrary overall phase factor.

In numerical calculations the Bloch states $\psi_{n\mathbf{k}}$ are computed on a regular mesh of k-points in the Brillouin zone; the integral in Eq. (1) is then replaced by a sum over the points in the mesh. In Ref. 16 an expression was derived for the gradient of the spread functional Ω with respect to an infinitesimal rotation $\delta U^{(\mathbf{k})}$ of the set of Bloch orbitals. The only information needed for calculating the gradient are the overlaps

$$M_{mn}^{(\mathbf{k},\mathbf{b})} = \langle u_{m\mathbf{k}} | u_{n,\mathbf{k}+\mathbf{b}} \rangle \,, \tag{5}$$

where $\mathbf{b}$ are vectors connecting a mesh point to its near neighbors. Once the gradient is computed, the minimization can proceed via a steepest-descent or conjugate-gradients algorithm.

In Ref. 16 the spread Ω was decomposed into two terms,

$$\Omega = \Omega_{\rm I} + \widetilde{\Omega}, \tag{6}$$

both of them non-negative. The first measures the k-space dispersion of the band projection operator, while the second reflects the extent to which the Wannier functions fail to be eigenfunctions of the band-projected position operators. $\Omega_{\rm I}$ will play a central role in the present work. For an isolated group of bands it is invariant under any gauge transformation (3), so that minimizing Ω amounts to minimizing $\widetilde{\Omega}$. When using a regular mesh of k-points, $\Omega_{\rm I}$ is given by

$$\Omega_{\rm I} = \frac{1}{N_{\rm kp}} \sum_{\mathbf{k},\mathbf{b}} w_b \sum_{m=1}^{N} \Big[1 - \sum_{n=1}^{N} \big| M_{mn}^{(\mathbf{k},\mathbf{b})} \big|^2 \Big], \tag{7}$$

where $N_{\rm kp}$ is the total number of k-points, N is the number of bands in the group, and w_b is a weight that arises from the discretization procedure by which derivatives with respect to $\mathbf{k}$ are approximated by finite differences.[16] The corresponding expression for $\widetilde{\Omega}$ can be found in Ref. 16.

III. MAXIMALLY-LOCALIZED WANNIER FUNCTIONS FOR ATTACHED BANDS

A. Description of the method

For definiteness let us suppose we want to "disentangle" the five d bands of copper from the s band which crosses them (see Fig. 1) and construct a set of well-localized WFs associated with the resulting d bands. Heuristically the d bands are the five narrow bands and the s band is the wide band. The difficulty arises because there are regions of k-space where all six bands are close together, so that as a result of hybridization "the distinction between d-band and s-band levels is not meaningful" (Ref. 1, p. 288).

Let us now outline our strategy, which can be divided in two steps. First we cut out an energy window that encompasses the N bands of interest ($N = 5$ in our example). Figs. 1(a) and 1(b) correspond to different choices for this energy window. At each k-point the number $N_{\mathbf{k}}$ of bands that fall inside the window is equal to or larger than the target number of bands N. This procedure defines an $N_{\mathbf{k}}$-dimensional Hilbert space $\mathcal{F}(\mathbf{k})$ spanned by the states $u_{n\mathbf{k}}$ within the window. If at some $\mathbf{k}$ $N_{\mathbf{k}} = N$, there is nothing to do there; if $N_{\mathbf{k}} > N$ our aim is to find

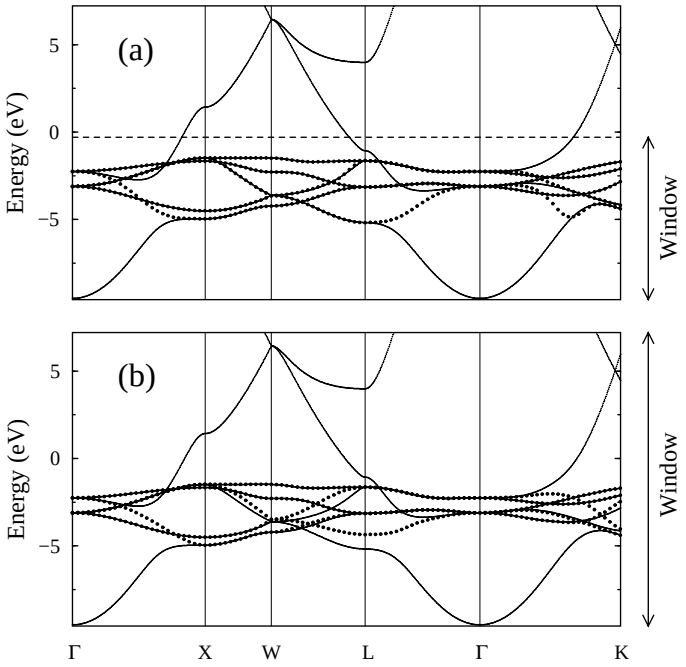

FIG. 1. Solid line: Calculated band structure of copper. Dotted line: Interpolated bands obtained from the five d-like Wannier functions. (a) and (b) differ in the choice of the energy window used to compute the Wannier functions ($[-9.59, -0.29]$ eV in (a) and $[-9.59, 7.21]$ eV in (b)). The zero of the energy scale is at the Fermi energy.

the N-dimensional subspace $\mathcal{S}(\mathbf{k}) \subseteq \mathcal{F}(\mathbf{k})$ that, among all possible N-dimensional subspaces of $\mathcal{F}(\mathbf{k})$, leads to the smallest $\Omega_{\rm I}$ (Eq. (7)). (Recall that for an isolated group of bands $\Omega_{\rm I}$ is gauge-invariant, since it is an intrinsic property of the manifold of states. Thus $\Omega_{\rm I}$ can be regarded as a functional of $\mathcal{S}(\mathbf{k})$.) In the second step we work within the optimal N-dimensional subspaces $\mathcal{S}(\mathbf{k})$ selected in the first step, and minimize $\widetilde{\Omega}$ using the algorithm of Marzari and Vanderbilt[16] summarized in the previous section. The end result is a set of N maximally localized WFs and the corresponding N energy bands. We emphasize that it is the first step (minimization of $\Omega_{\rm I}$) that is new with respect to Ref. 16.

B. Physical interpretation of $\Omega_{\rm I}$

Why is minimizing $\Omega_{\rm I}$ a sensible strategy for picking out the d-bands? This can be understood by noting that heuristically $\Omega_{\rm I}$ measures the "change of character" of the states across the Brillouin zone.[16] Indeed, Eqs. (5) and (7) show that $\Omega_{\rm I}$ is small whenever $|\langle u_{n\mathbf{k}} | u_{m,\mathbf{k}+\mathbf{b}} \rangle|^2$, the square of the magnitude of the overlap between states at nearby k-points, is large. Thus by minimizing $\Omega_{\rm I}$ we are choosing self-consistently at every $\mathbf{k}$ the subspace $\mathcal{S}(\mathbf{k})$ that has minimum "spillage" or mismatch (see below) as $\mathbf{k}$ is varied. In the present example this optimal "global smoothness of connection" will be achieved by

keeping the five well-localized d-like states and excluding the more delocalized s-like state. We will gain more intuition about the meaning of minimizing $\Omega_{\rm I}$ while discussing specific examples in Sec. IV.

What is meant by "spillage"[16,32] becomes clear once we rewrite Eq. (7) as

$$\Omega_{\rm I} = \frac{1}{N_{\rm kp}} \sum_{\mathbf{k},\mathbf{b}} w_b \, T_{\mathbf{k},\mathbf{b}} \tag{8}$$

with

$$T_{\mathbf{k},\mathbf{b}} = N - \sum_{m,n} \left| M_{mn}^{(\mathbf{k},\mathbf{b})} \right|^2 = \mathrm{tr}[\hat{P}_{\mathbf{k}} \hat{Q}_{\mathbf{k}+\mathbf{b}}], \tag{9}$$

where $\hat{P}_{\mathbf{k}} = \sum_n |u_{n\mathbf{k}}\rangle\langle u_{n\mathbf{k}}|$ is the projector onto $\mathcal{S}(\mathbf{k})$, $\hat{Q}_{\mathbf{k}} = \mathbf{1} - \hat{P}_{\mathbf{k}}$, and the band indices m, n run over $1, \ldots, N$. $T_{\mathbf{k},\mathbf{b}}$ is called the "spillage" between the spaces $\mathcal{S}(\mathbf{k})$ and $\mathcal{S}(\mathbf{k}+\mathbf{b})$ because it measures the degree of mismatch between them, vanishing when they are identical.

Further discussion of the geometrical and physical interpretation of $\Omega_{\rm I}$ can be found in Refs. 13 and 16. In particular, it has been shown that the value of $\Omega_{\rm I}$ associated with the valence bands of an insulator is the experimentally measurable mean-square quantum fluctuation of the ground state macroscopic polarization.[13] This can be interpreted as the quadratic spread of an appropriately defined collective center-of-mass distribution for the electrons, and can be recast as an electronic localization length squared. Hence our procedure of minimizing $\Omega_{\rm I}$ selects the N-dimensional subspaces $\mathcal{S}(\mathbf{k})$ where the electrons are most localized in the above sense (assuming for the purpose of this argument that all the electron states in those subspaces are occupied).

Finally we note in passing that our two-step procedure of minimizing first $\Omega_{\rm I}$ and then $\widetilde{\Omega}$ is in principle different from directly minimizing their sum Ω. In view of the discussion presented above, we believe that the procedure adopted here is conceptually the more natural of the two, although we would expect them to yield similar results in practice. Also, as we will now show, the separate minimization of $\Omega_{\rm I}$ turns out to be a particularly simple and robust procedure.

C. Iterative minimization of $\Omega_{\rm I}$

Since the functional (7) that we wish to minimize couples states at different k-points, the problem has to be solved self-consistently throughout the Brillouin zone. Our strategy is to proceed iteratively until the optimal "global smoothness of connection" is achieved. On the i-th iteration we go through all the k-points in the grid, and for each of them we find N orthonormal states $u_{n\mathbf{k}}^{(i)}$, defining a subspace $\mathcal{S}^{(i)}(\mathbf{k}) \subseteq \mathcal{F}(\mathbf{k})$ such that the "spillage" over the neighboring subspaces $\mathcal{S}^{(i-1)}(\mathbf{k}+\mathbf{b})$ from the previous iteration is as small as possible (Fig. 2).

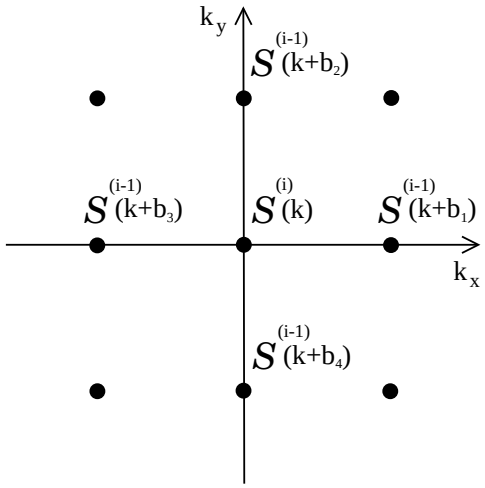

FIG. 2. Schematic representation of the subspaces of Bloch-like states on a grid of k-points. Our procedure consists of iteratively minimizing the "spillage", or degree of mismatch (see text), between the subspaces at neighboring k-points.

Using Lagrange multipliers to enforce orthonormality, the stationarity condition at the i-th iteration reads

$$\frac{\delta \Omega_{\rm I}^{(i)}}{\delta u_{m\mathbf{k}}^{(i)*}} + \sum_{n=1}^{N} \Lambda_{nm,\mathbf{k}}^{(i)} \frac{\delta}{\delta u_{m\mathbf{k}}^{(i)*}} \left[\left\langle u_{m\mathbf{k}}^{(i)} \middle| u_{n\mathbf{k}}^{(i)} \right\rangle - \delta_{m,n} \right] = 0, \tag{10}$$

where $\Lambda_{\mathbf{k}}^{(i)}$ is an $N \times N$ matrix. Let

$$\Omega_{\rm I}^{(i)} = \frac{1}{N_{\rm kp}} \sum_{\mathbf{k}=1}^{N_{\rm kp}} \omega_{\rm I}^{(i)}(\mathbf{k}) \tag{11}$$

where, according to Eq. (8),

$$\begin{aligned} \omega_{\rm I}^{(i)}(\mathbf{k}) &= \sum_{\mathbf{b}} w_b \, T_{\mathbf{k},\mathbf{b}}^{(i)} \\ &= \sum_{\mathbf{b}} w_b \sum_{m=1}^{N} \left[1 - \sum_{n=1}^{N} \left| \left\langle u_{m\mathbf{k}}^{(i)} \middle| u_{n,\mathbf{k}+\mathbf{b}}^{(i-1)} \right\rangle \right|^2 \right]. \end{aligned} \tag{12}$$

The first term in Eq. (10) now becomes

$$\frac{\delta \Omega_{\rm I}^{(i)}}{\delta u_{m\mathbf{k}}^{(i)*}} = \frac{1}{N_{\rm kp}} \left\{ \frac{\delta \omega_{\rm I}^{(i)}(\mathbf{k})}{\delta u_{m\mathbf{k}}^{(i)*}} + \sum_{\mathbf{b}} \frac{\delta \omega_{\rm I}^{(i)}(\mathbf{k}+\mathbf{b})}{\delta u_{m\mathbf{k}}^{(i)*}} \right\}. \tag{13}$$

From Eq. (12) we find

$$\frac{\delta \omega_{\rm I}^{(i)}(\mathbf{k})}{\delta u_{m\mathbf{k}}^{(i)*}} = -\sum_{\mathbf{b}} w_b \, \hat{P}_{\mathbf{k}+\mathbf{b}}^{(i-1)} \left| u_{m\mathbf{k}}^{(i)} \right\rangle, \tag{14}$$

where $\hat{P}_{\mathbf{k}+\mathbf{b}}^{(i-1)}$ is the projector onto $\mathcal{S}^{(i-1)}(\mathbf{k}+\mathbf{b})$. Likewise, one easily obtains

$$\frac{\delta \omega_{\rm I}^{(i)}(\mathbf{k}+\mathbf{b})}{\delta u_{m\mathbf{k}}^{(i)*}} = -w_b \, \hat{P}_{\mathbf{k}+\mathbf{b}}^{(i-1)} \left| u_{m\mathbf{k}}^{(i)} \right\rangle. \tag{15}$$

Combining the previous equations, the stationarity condition (10) becomes

$$\Big[\sum_{\mathbf{b}} w_b \hat{P}^{(i-1)}_{\mathbf{k}+\mathbf{b}}\Big] \Big| u^{(i)}_{m\mathbf{k}} \Big\rangle = \sum_{n=1}^{N} \widetilde{\Lambda}^{(i)}_{nm,\mathbf{k}} \Big| u^{(i)}_{n\mathbf{k}} \Big\rangle \,, \qquad (16)$$

where $\widetilde{\Lambda}^{(i)}_{nm,\mathbf{k}} = (N_{\rm kp}/2)\Lambda^{(i)}_{nm,\mathbf{k}}$. By choosing a unitary transformation that diagonalizes $\widetilde{\Lambda}^{(i)}_{\mathbf{k}}$, this can be recast as an eigenvalue equation:

$$\Big[\sum_{\mathbf{b}} w_b \hat{P}^{(i-1)}_{\mathbf{k}+\mathbf{b}}\Big] \Big| u^{(i)}_{m\mathbf{k}} \Big\rangle = \lambda^{(i)}_{m\mathbf{k}} \Big| u^{(i)}_{m\mathbf{k}} \Big\rangle \,. \qquad (17)$$

The eigenvalues of the above equation obey $0 \le \lambda^{(i)}_{m\mathbf{k}} \le \sum_{\mathbf{b}} w_b$; in particular, $\lambda^{(i)}_{m\mathbf{k}} < \sum_{\mathbf{b}} w_b$ whenever the eigenstate $u^{(i)}_{m\mathbf{k}}$ does not lie completely within all of the nearby subspaces $\mathcal{S}^{(i-1)}(\mathbf{k}+\mathbf{b})$. Combining Eqs. (12) and (17), we find

$$\omega^{(i)}_{\rm I}(\mathbf{k}) = N \sum_{\mathbf{b}} w_b - \sum_{m=1}^{N} \lambda^{(i)}_{m\mathbf{k}}. \qquad (18)$$

It is clear from Eqs. (11) and (18) that when constructing $\mathcal{S}^{(i)}(\mathbf{k})$ one should pick the N eigenvectors of Eq. (17) with largest eigenvalues, so as to ensure that the stationary point corresponds to the absolute minimum of $\Omega^{(i)}_{\rm I}$.

Self-consistency is achieved when $\mathcal{S}^{(i)}(\mathbf{k}) = \mathcal{S}^{(i-1)}(\mathbf{k})$ at all the grid points. We have encountered cases where the iterative procedure outlined above was not stable. In those cases, the problem was solved by using as the input for the present step a linear mixing of the input and output subspaces from the previous step. More precisely, the eigenvalue equation (17) was replaced by

$$\Big\{\sum_{\mathbf{b}} w_b \Big[\hat{\mathcal{P}}^{(i)}_{\mathbf{k}+\mathbf{b}}\Big]_{\rm in}\Big\} \Big| u^{(i)}_{m\mathbf{k}} \Big\rangle = \lambda^{(i)}_{m\mathbf{k}} \Big| u^{(i)}_{m\mathbf{k}} \Big\rangle \,, \qquad (19)$$

where

$$\Big[\hat{\mathcal{P}}^{(i)}_{\mathbf{k}+\mathbf{b}}\Big]_{\rm in} = \alpha \hat{P}^{(i-1)}_{\mathbf{k}+\mathbf{b}} + (1-\alpha)\Big[\hat{\mathcal{P}}^{(i-1)}_{\mathbf{k}+\mathbf{b}}\Big]_{\rm in} \qquad (20)$$

with $0 < \alpha \le 1$.[33] A typical value is α=0.5.

In practice we solve Eq. (19) in the basis of the original $N_{\mathbf{k}}$ Bloch eigenstates $u_{n\mathbf{k}}$ inside the energy window. Each iteration then amounts to diagonalizing the following $N_{\mathbf{k}} \times N_{\mathbf{k}}$ Hermitian matrix at every $\mathbf{k}$:

$$Z^{(i)}_{mn}(\mathbf{k}) = \Big\langle u_{m\mathbf{k}} \Big| \sum_{\mathbf{b}} w_b \Big[\hat{\mathcal{P}}^{(i)}_{\mathbf{k}+\mathbf{b}}\Big]_{\rm in} \Big| u_{n\mathbf{k}} \Big\rangle. \qquad (21)$$

Since these are small matrices, each step of the iterative procedure is computationally cheap. In particular, the time-consuming computation of the overlap matrices $M^{(\mathbf{k},\mathbf{b})}$ of Eq. (5) can be done once and for all at the beginning, using the original Bloch eigenstates inside the energy window; their subsequent update during the iterative minimization is very inexpensive. An analogous situation occurs when updating the matrices $U^{(\mathbf{k})}$ in Eq. (3) during the minimization of $\widetilde{\Omega}$ to obtain the "maxloc" WFs.[16]

D. Initial guess for the subspaces

In order to start the iterative minimization of $\Omega_{\rm I}$, the user should provide an initial guess for the subspaces $\mathcal{S}(\mathbf{k})$. We have found that the minimization procedure is quite robust, in the sense that it is able to arrive at the global minimum starting from a very rough initial guess. In practice we usually select the initial subspaces following a strategy very similar to the one outlined in Ref. 16 for starting the minimization of $\widetilde{\Omega}$.

A set of N localized trial orbitals $g_n(\mathbf{r})$ is chosen corresponding to some rough initial guess at the WFs, and these are then projected onto the $N_{\mathbf{k}}$ Bloch eigenstates inside the energy window,

$$|\,\phi_{n\mathbf{k}}\rangle = \sum_{m=1}^{N_{\mathbf{k}}} A_{mn}\,|\,\psi_{m\mathbf{k}}\rangle \,, \qquad (22)$$

where $A_{mn} = \langle \psi_{m\mathbf{k}} | g_n \rangle$ is an $N_{\mathbf{k}} \times N$ matrix. The resulting N orbitals are then orthonormalized via Löwdin's symmetric orthogonalization procedure,[34] i.e.,

$$\begin{aligned} \Big|\,\psi^{(0)}_{n\mathbf{k}}\Big\rangle &= \sum_{m=1}^{N} (S^{-1/2})_{mn}\,|\,\phi_{m\mathbf{k}}\rangle \\ &= \sum_{m=1}^{N_{\mathbf{k}}} (AS^{-1/2})_{mn}\,|\,\psi_{m\mathbf{k}}\rangle \,, \end{aligned} \qquad (23)$$

where $S_{mn} = \langle \phi_{m\mathbf{k}} | \phi_{n\mathbf{k}} \rangle = (A^\dagger A)_{mn}$. Finally these Bloch-like functions are converted to cell-periodic functions $u^{(0)}_{n\mathbf{k}} = e^{-i\mathbf{k}\cdot\mathbf{r}}\psi^{(0)}_{n\mathbf{k}}$. The matrix $AS^{-1/2}$ can easily be computed by performing the singular-value decomposition $A = ZDV$,[35] where Z and V are $N_{\mathbf{k}} \times N_{\mathbf{k}}$ and $N \times N$ unitary matrices respectively, and D is $N_{\mathbf{k}} \times N$ and diagonal. This leads to $AS^{-1/2} = Z\mathbf{1}V$, where $\mathbf{1}$ is the $N_{\mathbf{k}} \times N$ identity matrix.

E. Minimization of $\widetilde{\Omega}$

At the end of the first step of our procedure (minimization of $\Omega_{\rm I}$) we are left at each k-point with an N-dimensional subspace $\mathcal{S}(\mathbf{k})$, and for definiteness we diagonalize the Hamiltonian inside this subspace to obtain N Bloch-like eigenfunctions $\widetilde{\psi}_{n\mathbf{k}} = e^{i\mathbf{k}\cdot\mathbf{r}}\widetilde{u}_{n\mathbf{k}}$ and eigenvalues $\widetilde{\epsilon}_{n\mathbf{k}}$. The second step is to find the $N \times N$ unitary matrices $U^{(\mathbf{k})}$ (Eq. (3)) that, applied to the $\widetilde{\psi}_{n\mathbf{k}}$, produce the rotated set of Bloch-like states that is transformed via (1) into the maximally-localized WFs $w_{n\mathbf{R}}$. This is done using the method of Marzari and Vanderbilt[16] for minimizing $\widetilde{\Omega}$, briefly discussed in Sec. II. An initial guess for the unitary matrices $U^{(\mathbf{k})}$ is obtained by projecting a set of N localized orbitals onto the states $\widetilde{\psi}_{n\mathbf{k}}$. Typically the same set of orbitals is used as in the initialization step for the minimization of $\Omega_{\rm I}$. (In our experience, when a

particularly bad choice of trial orbitals is made, the minimization of $\Omega_{\rm I}$ is less likely to become trapped in local minima than the minimization of $\widetilde{\Omega}$.)

F. Interpolated band structure

Starting from the "maxloc" WFs, the corresponding energy bands can be computed at arbitrary points in the Brillouin zone using a Slater-Koster interpolation scheme.[27,31,36] Of course, the interpolation could proceed directly from the non-rotated states $\widetilde{u}_{n\mathbf{k}}$; however, use of the optimally rotated ones ensures that the interpolated band structure is as smooth as possible.[37]

The interpolation procedure involves first calculating the Hamiltonian matrix for the rotated states,

$$H^{(\rm rot)}(\mathbf{k}) = \left(U^{(\mathbf{k})}\right)^\dagger \widetilde{H}(\mathbf{k})\, U^{(\mathbf{k})}, \tag{24}$$

where $\widetilde{H}_{mn}(\mathbf{k}) = \widetilde{\epsilon}_{m\mathbf{k}}\delta_{m,n}$. Next we Fourier transform $H^{(\rm rot)}(\mathbf{k})$ into a set of $N_{\rm kp}$ Bravais lattice vectors $\mathbf{R}$ within a Wigner-Seitz supercell centered around $\mathbf{R}=0$:

$$\begin{aligned} H^{(\rm rot)}_{mn}(\mathbf{R}) &= \Big(\sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{R}}\, H^{(\rm rot)}_{mn}(\mathbf{k})\Big)/N_{\rm kp} \\ &= \big\langle w_{m\mathbf{0}}\big|\hat{H}\big|w_{n\mathbf{R}}\big\rangle, \end{aligned} \tag{25}$$

where $\hat{H}$ is the effective one-particle Hamiltonian. Finally we Fourier transform back to an arbitrary k-point,

$$H^{(\rm rot)}_{mn}(\mathbf{k}') = \sum_{\mathbf{R}} e^{i\mathbf{k}'\cdot\mathbf{R}}\, H_{mn}(\mathbf{R}), \tag{26}$$

and diagonalize the resulting matrix to find the interpolated energy eigenvalues.

G. Inner energy window

In some situations one wants to construct orbitals that describe the original bands *exactly* only in a limited energy range. This can occur when studying transport properties for which only the states within some small energy range of the Fermi level (say, ± 1 eV) are relevant. The challenge is to construct orbitals that achieve that goal while remaining as localized as possible. What the resulting interpolated bands look like outside the energy range of interest is largely immaterial, since it will not affect the low-energy physics. (Typically they will tend to remain close in energy to the target range of interest.[23])

A simple extension of the formalism described in the previous sections can produce such orbitals. The idea is to introduce a second ("inner") energy window – contained within our original ("outer") window – inside which the original bands are to be described exactly. Let $M_{\mathbf{k}}$ be the number of bands that fall within the inner window at $\mathbf{k}$, so that $M_{\mathbf{k}} \le N \le N_{\mathbf{k}}$. Then we have to minimize $\Omega_{\rm I}$ under the constraint that the $M_{\mathbf{k}}$ original Bloch states inside the inner window must be included in the subspace $\mathcal{S}(\mathbf{k})$. We are therefore only free to choose the remaining $N-M_{\mathbf{k}}$ states when constructing $\mathcal{S}(\mathbf{k})$. Those will have to be extracted from the subspace spanned by the $N_{\mathbf{k}} - M_{\mathbf{k}}$ original Bloch eigenstates that are inside the outer window but outside the inner window. That can be achieved by a straightforward modification of the iterative procedure described in Sec. III C: The matrix $Z^{(i)}(\mathbf{k})$ in Eq. (21) becomes an $(N_{\mathbf{k}} - M_{\mathbf{k}}) \times (N_{\mathbf{k}} - M_{\mathbf{k}})$ matrix, and we pick the $N - M_{\mathbf{k}}$ leading eigenvectors.

The only remaining issue is how to modify the initialization procedure of Sec. III D in order to accommodate the inner window. Since the first $M_{\mathbf{k}}$ basis vectors of the trial subspaces $\mathcal{S}(\mathbf{k})$ are predetermined, we want the modified procedure to provide the remaining $N-M_{\mathbf{k}}$ vectors. Let $\mathcal{G}(\mathbf{k})$ be an N-dimensional space obtained by projecting the N trial orbitals onto the $N_{\mathbf{k}}$ states inside the outer window, as described in Sec. III D. Let $P_{\mathcal{G}}(\mathbf{k})$ be the $N_{\mathbf{k}} \times N_{\mathbf{k}}$ matrix that is the projection operator onto $\mathcal{G}(\mathbf{k})$ as expressed in the space $\mathcal{F}(\mathbf{k})$. Similarly, define $P_{\rm inner}(\mathbf{k})$ as the $N_{\mathbf{k}} \times N_{\mathbf{k}}$ projection matrix onto the inner window states, and $Q_{\rm inner}(\mathbf{k}) = \mathbf{1} - P_{\rm inner}(\mathbf{k})$. Then choose the remaining $N - M_{\mathbf{k}}$ basis vectors to be the eigenvectors corresponding to the $N - M_{\mathbf{k}}$ largest eigenvalues of

$$Q_{\rm inner}(\mathbf{k}) P_{\mathcal{G}}(\mathbf{k}) Q_{\rm inner}(\mathbf{k}) |v\rangle = \lambda |v\rangle. \tag{27}$$

Such vectors have the desired properties: (i) They are orthogonal to the states inside the inner window, and (ii) because $\lambda = \langle v|P_{\mathcal{G}}(\mathbf{k})|v\rangle$, it is clear that by choosing the eigenvectors with the largest eigenvalues we guarantee that their overlap with the space $\mathcal{G}(\mathbf{k})$ is as large as possible, while satisfying the constraint (i).

Other kinds of constraints on the minimization of $\Omega_{\rm I}$ may also be useful. For instance, one might want to "pin down" the desired bands at high-symmetry k-points to ensure that the interpolated bands coincide with them at those points.

IV. RESULTS

A. Computational details

The calculations were performed within the local-density approximation to density-functional theory, using a plane-wave basis set and Troullier-Martins norm-conserving pseudopotentials[38] in the Kleinman-Bylander representation. The energy cutoff was set to 75 Ry for copper and 35 Ry for silicon, and the lattice constants were 6.822 bohr and 10.260 bohr respectively. The computed self-consistent Bloch eigenfunctions and eigenvalues that fell inside the prescribed energy window were stored to disk. They were used as the input for the minimization of $\Omega_{\rm I}$, which was carried out as a separate,

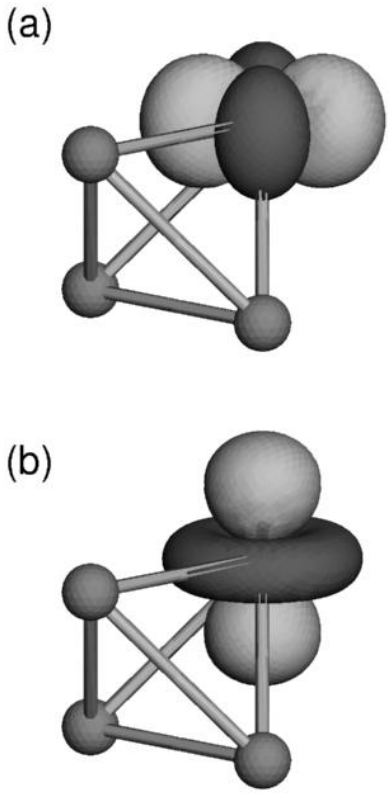

FIG. 3. Contour-surface plots of the two e_g Wannier functions associated with the "disentangled" d bands of copper shown in Fig. 1(b). The amplitudes are $+0.5/\sqrt{v}$ (light gray) and $-0.5/\sqrt{v}$ (dark gray), where v is the volume of the primitive cell.

postprocessing operation. This produced an optimal subspace characterized by a new set of N Bloch eigenfunctions and eigenvalues per k-point, which were taken as the input for constructing the "maxloc" WFs and the interpolated bands. In all the cases we have found the "maxloc" WFs to be real (apart from an overall phase factor), as was already the case when dealing with isolated groups of bands.[16] The self-consistent calculations were performed on a $10 \times 10 \times 10$ Monkhorst-Pack mesh of k-points for copper, and $6 \times 6 \times 6$ for silicon. During the minimization of $\Omega_{\rm I}$ and $\widetilde{\Omega}$ a $10 \times 10 \times 10$ uniform grid was used for both copper and silicon. This grid was shifted in order to include the Γ point ($k = 0$), so as to ensure that the "maxloc" WFs have the desired symmetry properties among themselves. (For instance, if a grid is used for silicon that does not include Γ, the four antibonding WFs in a unit cell do not all have the same spread.) The mixing parameter α in Eq. (20) was set to 0.5.

B. Copper

Wannier functions for noble and transition metals have previously been computed using various approaches.[24–27] Below, taking copper as an example, we show how the present scheme can be used to "disentangle" the narrow d bands from the nearly-free-electron bands, allowing us to treat each group of WFs separately. Alternatively, one can also treat the narrow and the nearly-free-electron bands as a single group.

1. Narrow d bands

First, an energy window was chosen such that at each k-point in the grid it contained six or seven energy eigenvalues. As indicated in Fig. 1, the precise range of the window is largely at our disposal; unless explicitly stated otherwise, the numbers given below pertain to Fig. 1(b). In order to extract the five d bands, we set $N = 5$ and initialized the minimization of both $\Omega_{\rm I}$ and $\widetilde{\Omega}$ from five trial Gaussians of r.m.s. width 1 bohr, each modulated by a different $l = 2$ angular eigenfunction. After ~ 50 iterative steps $\Omega_{\rm I}$ was fully converged, having decreased from an initial value of 9.957 bohr^2 to 8.483 bohr^2. During the subsequent minimization of $\widetilde{\Omega}$ the total Wannier spread Ω decreased only slightly, from 8.563 bohr^2 to 8.556 bohr^2. In agreement with previous experience on isolated groups of bands,[16] we found for the d bands that at the minimum $\Omega_{\rm I} \gg \widetilde{\Omega}$.

The bands obtained by interpolation using the five "maxloc" WFs are shown as dotted lines in Fig. 1, together with the original band structure. As expected, whenever the dispersive s-like band is far from the narrow d bands, so that they retain their separate identities, the interpolated bands are very close to the narrow bands. However, whenever the six bands are close together, and thus strongly hybridized, the interpolated bands remain narrow, which suggests that they are mainly d-like in character. (Heuristically they can be viewed as the bands obtained by artificially "switching off" the Hamiltonian matrix elements between s and d WFs, i.e., by removing the hybridization.) The d character is confirmed by inspection of the contour-surface plots of the "maxloc" WFs, two of which are shown in Fig. 3. The quadratic spreads of the five WFs are not exactly equal, because of the $e_g - t_{2g}$ splitting of the d-states; those shown in Fig. 3 (e_g orbitals) have a spread of 1.700 bohr^2 each, whereas the remaining three (t_{2g} orbitals) each have a spread of 1.718 bohr^2. These numbers are only slightly

TABLE I. Variation of the optimal Wannier spread Ω and its gauge-invariant part $\Omega_{\rm I}$ (in bohr^2) with the choice of energy window range (in eV), for the d bands of copper.

Window range		Total spread	
Min	Max	$\Omega_{\rm I}$	Ω
−9.59	−0.29	15.373	16.489
−9.59	2.21	10.404	10.621
−9.59	7.21	8.483	8.556
−9.59	12.21	7.634	7.667

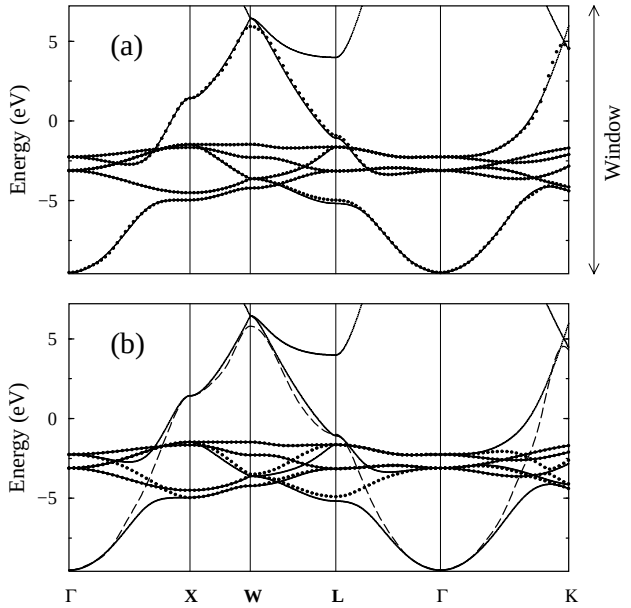

FIG. 4. (a) Dotted lines: the s-d bands of copper obtained by extracting the optimal six-dimensional subspace $\mathcal{S}_6(\mathbf{k})$ inside the window. (b) Dotted lines: d bands associated with optimal five-dimensional subspace $\mathcal{S}_5(\mathbf{k}) \subset \mathcal{S}_6(\mathbf{k})$. Dashed line: s band $\mathcal{S}_1(\mathbf{k})$ isolated by taking the complement of $\mathcal{S}_5(\mathbf{k})$.

larger than the ones reported in Table III of Ref. 27, obtained using a different method and a sparser sampling of the Brillouin zone.

In our procedure there is one adjustable parameter, namely the range of the energy window. This range should be wide enough that it encompasses the bands of interest, but not be so wide that it also includes other bands of similar character (e.g., higher d bands). In the limit of a very wide window the spaces $\mathcal{F}(\mathbf{k})$ would contain a complete set of states, so that by mixing in states far away from the energy range of interest but of similar character, the spread of the WFs could be made arbitrarily small (and the corresponding bands would become flat). Table I shows how the optimal Wannier spreads are affected by varying the window range within reasonable bounds. As anticipated, the spread decreases with increasing energy range.[39] The change in the interpolated energy bands is less pronounced, although they do become somewhat narrower (compare Figs. 1(a) and 1(b)). In particular, the upward shift of the lowest interpolated band at L is caused by mixing with the seventh band, which has the same symmetry label (L_1).[40]

2. Nearly-free-electron band

The unconstrained minimization of $\Omega_{\rm I}$ usually produces narrow bands, since the character of the Bloch states in such bands tends to have only a small variation across the Brillouin zone, corresponding to well-localized electrons (this may not be the case in the presence of avoided crossings). The method is therefore ideally suited for directly extracting the narrow d bands from the s-d complex. If instead one is interested in isolating the wider, nearly-free-electron s band, direct minimization of $\Omega_{\rm I}$ for one-dimensional subspaces is not the appropriate strategy. Instead one can proceed as follows. First choose an energy window that includes the s-d band complex (we used the one indicated in Fig. 1(b)). Then minimize $\Omega_{\rm I}$ with $N = 6$; this produces a six-dimensional subspace $\mathcal{S}_6(\mathbf{k})$ throughout the Brillouin zone that consists of the s–d band complex. Next extract the five d bands by minimizing $\Omega_{\rm I}$ within $\mathcal{S}_6(\mathbf{k})$ choosing $N = 5$; this yields a space $\mathcal{S}_5(\mathbf{k}) \subset \mathcal{S}_6(\mathbf{k})$. The difference between the two is a one-dimensional space $\mathcal{S}_1(\mathbf{k})$ containing the desired band. Fig. 4(a) shows the bands associated with $\mathcal{S}_6(\mathbf{k})$, and Fig. 4(b) shows the bands corresponding to $\mathcal{S}_5(\mathbf{k})$ and $\mathcal{S}_1(\mathbf{k})$.

TABLE II. Spreads of the "maxloc" WFs for the separate d-band and s-band subspaces ($\mathcal{S}_5$ and $\mathcal{S}_1$), and for the combined s-d subspace $\mathcal{S}_6$. The numbers in parentheses are the $\Omega_{\rm I}$ values, and t stands for tetrahedral-interstitial-centered orbital. The corresponding bands are displayed in Fig. 4.

Two separate subspaces			One combined subspace		
d_{e_g}	1.710		d_{e_g}	1.731	
d_{e_g}	1.710		d_{e_g}	1.731	
$d_{t_{2g}}$	1.808		$d_{t_{2g}}$	2.328	
$d_{t_{2g}}$	1.808		$d_{t_{2g}}$	2.328	
$d_{t_{2g}}$	1.808		$d_{t_{2g}}$	2.254	
$\Omega_{\rm min}[\mathcal{S}_5]$	8.844	(8.745)			
t	12.929		t	10.263	
$\Omega_{\rm min}[\mathcal{S}_1]$	12.929	(10.826)	$\Omega_{\rm min}[\mathcal{S}_6]$	20.634	(16.506)

In Table II are presented the optimal Wannier spreads for the different subspaces. We find that the spread of the s-like WF is considerably smaller than the $\sim$45 bohr^2 reported in Table III of Ref. 27. Moreover, contrary to what one might have expected, that WF is centered not on an atom, but on a tetrahedral interstitial site, as shown in Fig. 5(a). Since there are two such sites per atom, a breaking of symmetry must have occurred when selecting the subspace $\mathcal{S}_6(\mathbf{k})$. Indeed there are two degenerate minima of $\Omega_{\rm I}$ with $N = 6$, one for each of the interstitial sites. If the minimization is initialized by projecting five d-like orbitals plus one s-like orbital, all atom-centered, the breaking of symmetry occurs spontaneously during the iterative procedure (the minimization of $\Omega_{\rm I}$ reaches a plateau, presumably a saddle point, and eventually the algorithm finds its way towards one of the two minima). If instead the s trial orbital is centered around one of the tetrahedral interstitial sites, the minimization starts inside the basin of attraction of the corresponding minimum.

Finally, as a simple illustration of the "inner window" idea of Sec. III G, we show in Fig. 6 the single band (N=1) that results when an inner window is selected in

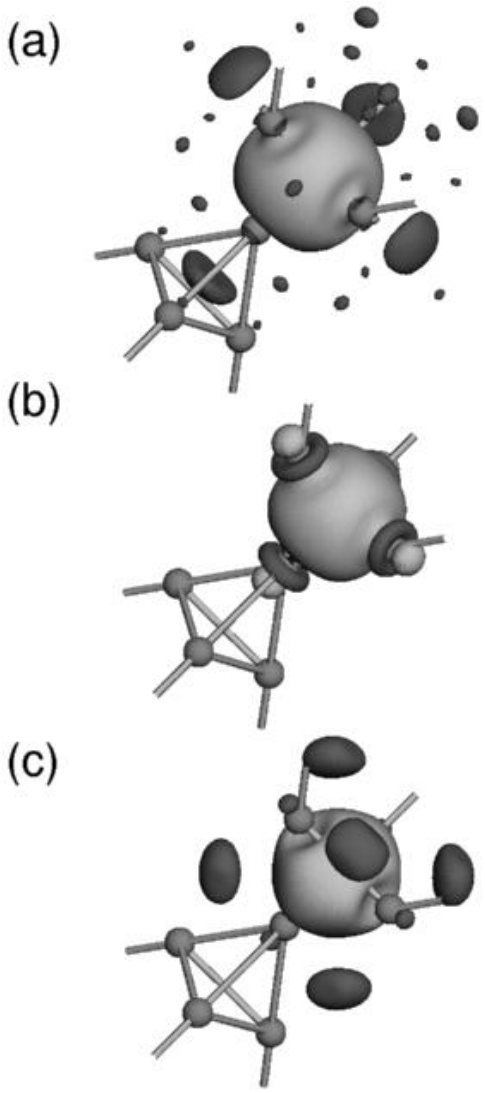

FIG. 5. Contour-surface plots of interstitial-centered "maxloc" WFs. (a) t-like WF associated with the subspace $\mathcal{S}_6(\mathbf{k})$ of Fig. 4 and Table II; (b) WF associated with the band in Fig. 6; (c) t-like WF associated with the subspace $\mathcal{S}_7(\mathbf{k})$ in Fig. 7(a) and Table III. The amplitudes are $+0.5/\sqrt{v}$ (light gray) and $-0.17/\sqrt{v}$, $-0.3/\sqrt{v}$, and $-0.25/\sqrt{v}$ (dark gray) in (a), (b), and (c) respectively.

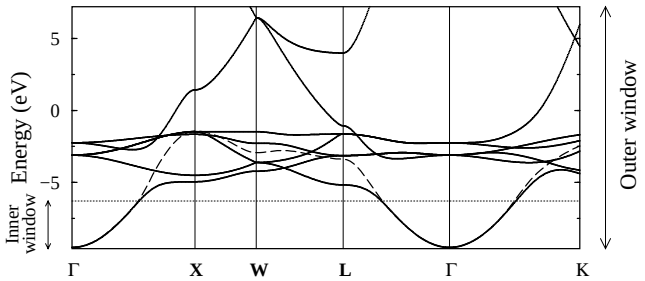

FIG. 6. Dashed line: Band obtained using both an inner and an outer energy window.

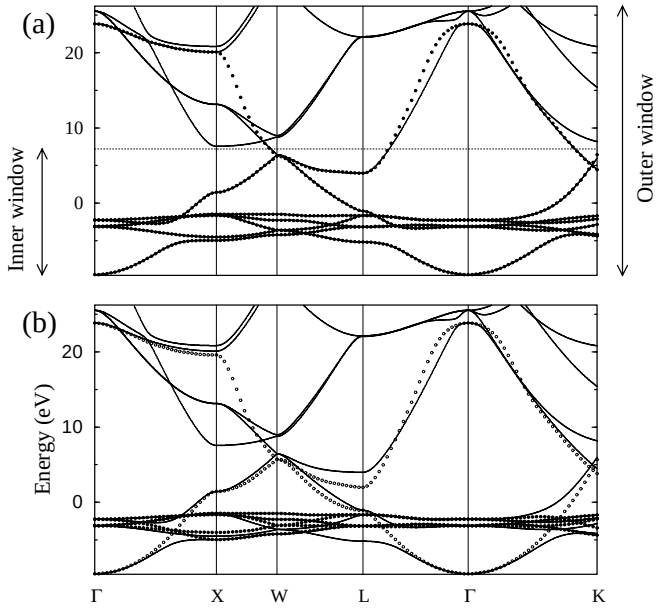

FIG. 7. (a) Dotted lines: Interpolated bands associated with the optimal subspace $\mathcal{S}_7(\mathbf{k})$ containing five d-like WFs and two tetrahedral-interstitial-centered WFs. (b) Dark dotted lines: d bands associated with optimal five-dimensional subspace $\mathcal{S}_5^{'}(\mathbf{k}) \subset \mathcal{S}_7(\mathbf{k})$. Light dotted lines: dispersive bands $S_2(\mathbf{k})$ isolated by taking the complement of $\mathcal{S}_5^{'}(\mathbf{k})$.

the energy range below the d bands. As expected, the interpolated band is identical to the original one inside that window. Moreover, it remains quite narrow outside, where it acquires a pronounced d character. (This means that the cost in $\Omega_{\rm I}$ of changing from s to d character is more than compensated by the smaller dispersion – and hence smaller $\Omega_{\rm I}$ – of the more localized d-like states.) Accordingly, the "maxloc" WF, shown in Fig. 5(b), is again centered at a tetrahedral interstitial site, like the WF of Fig. 5(a), but now it has a substantial admixture of d-like satellites and a smaller spread, $\Omega = 7.323$ bohr^2 ($\Omega_{\rm I} = 7.306$ bohr^2).

The results of this Section indicate that the occurrence of a symmetry breaking in the minimization of $\Omega_{\rm I}$ with a "maxloc" WF centered at a tetrahedral interstitial site appears to be a rather robust result.

3. Symmetric two-WF description of dispersive bands

Remarkably, we find that the symmetry can be restored, and a more faithful overall description of the bands can be achieved, by bringing in just one more dispersive band and working with a *set of seven WFs*. More precisely, we choose an energy window such as the one indicated in Fig. 7(a), containing seven or more bands, and minimize $\Omega_{\rm I}$ with $N = 7$. (To ensure that the low-energy part of the band complex is well described, we freeze it inside an inner window.) After applying the localization procedure we obtain, besides the five d orbitals, *two*

TABLE III. Spreads of the "maxloc" WFs for the separate d-band and low-lying dispersive bands subspaces ($\mathcal{S}_5^{'}$ and $\mathcal{S}_2$), and for the combined subspace $\mathcal{S}_7$. The numbers in parentheses are the $\Omega_{\rm I}$ values, and t stands for tetrahedral-interstitial-centered orbital. The corresponding bands are displayed in Fig. 7.

Two separate subspaces			One combined subspace		
d_{e_g}	1.687		d_{e_g}	1.687	
d_{e_g}	1.686		d_{e_g}	1.687	
$d_{t_{2g}}$	1.472		$d_{t_{2g}}$	1.737	
$d_{t_{2g}}$	1.472		$d_{t_{2g}}$	1.737	
$d_{t_{2g}}$	1.472		$d_{t_{2g}}$	1.737	
$\Omega_{\rm min}[\mathcal{S}_5^{'}]$	7.788	(7.751)			
t	8.568		t	7.812	
t	8.568		t	7.812	
$\Omega_{\rm min}[\mathcal{S}_2]$	17.136	(16.822)	$\Omega_{\rm min}[\mathcal{S}_7]$	24.209	(22.034)

equivalent WFs, each centered at one of the two tetrahedral interstitial sites. One of the latter is shown in Fig. 5(c). The optimal Wannier spreads are given in Table III; it can be seen that the spread of each of the two interstitial WFs is considerably smaller than that of the single interstitial WF in Table II and Fig. 5(a).

Fig. 7(b) shows the d-like bands associated with the optimal five-dimensional subspace $\mathcal{S}_5^{'}(\mathbf{k}) \subset \mathcal{S}_7(\mathbf{k})$, as well as the dispersive bands associated with $S_2(\mathbf{k})$, the complement of $\mathcal{S}_5^{'}(\mathbf{k})$ inside $\mathcal{S}_7(\mathbf{k})$. There is an upward shift in energy of the states X_3, W_3, and L_1 in the narrow bands, due to mixing with the states of the same symmetry in the dispersive bands, which suffer a downward shift of the same magnitude.

The fact that our procedure naturally generates a pair of WFs centered at the tetrahedral interstitial sites can be rationalized in terms of a tight-binding description of the nearly-free electron states. The tetrahedral interstitial sites form a simple cubic lattice, so that in view of Fig. 5(c) one might imagine that the electronic states of these WFs would be roughly analogous to those of a nearest-neighbor tight-binding model of s orbitals on the sites of a simple cubic lattice. Indeed we have checked that the main qualitative features of the interpolated bands associated with the two interstitial-centered WFs (light dotted lines in Fig. 7(b)) are captured by such a tight-binding model, but folded back into the fcc Brillouin zone to give two bands instead of one.

The quality of the interpolated bands in Fig. 7(a) suggests that the two tetrahedral-interstitial-centered orbitals (which we denote as t orbitals) complement the five atom-based d orbitals nicely to form a basis (t^2d^5) for a tight-binding parametrization of the copper bands. This requires only one more basis function than the traditional "minimal basis"[41] sd^5 (five d plus one s atomic orbitals), while still remaining more economical than the sp^3d^5 basis.[42] The three bases are compared in Table IV. At each high-symmetry k-point we list, in order of increasing energy, the symmetry labels of the states that occur in a detailed band-structure calculation (e.g., Ref. 40), and then whether or not they are captured by each of the tight-binding bases. Inspection of the table clarifies that the t^2d^5 basis has some very attractive features. Whereas the sd^5 basis misses the $X_{4'}$ state[41] (unoccupied p-like state not far above E_F) and, even more importantly, the $L_{2'}$ state (occupied p-like state just below E_F), t^2d^5 gets the symmetries right up to at least the first state above E_F at each high-symmetry k-point. Even sp^3d^5 does not do this, failing at the Γ point, since the state $\Gamma_{2'}$ has f character. A consequence of this analysis is that the t orbitals cannot be constructed solely from s and p orbitals. This can also be seen from Fig. 5(c): The positive-amplitude central portion of the WF can be interpreted in terms of a superposition of four sp hybrids coming from each of the four surrounding copper atoms and pointing towards the interstitial; however this picture cannot account for the six negative lobes.

To conclude, we note that the sp^3d^5 description can also be obtained from our procedure, by minimizing $\Omega_{\rm I}$ with $N = 9$ within a window containing eleven or more bands (e.g., with the upper bound at 32.2 eV). The "maxloc" WFs are then five atom-centered d-like orbitals

TABLE IV. A list, in order of increasing energy, of the symmetry labels of selected states in the band structure of copper (taken from Ref. 40), and whether or not they are captured by each of the tight-binding bases discussed in the text. An asterisk (*) indicates that the state is occupied.

	Degeneracy	sd^5	t^2d^5	sp^3d^5
Γ_1	1^*	yes	yes	yes
$\Gamma_{25'}$	3^*	yes	yes	yes
Γ_{12}	2^*	yes	yes	yes
$\Gamma_{2'}$	1	–	yes	–
Γ_{15}	3	–	–	yes
X_1	1^*	yes	yes	yes
X_3	1^*	yes	yes	yes
X_2	1^*	yes	yes	yes
X_5	2^*	yes	yes	yes
$X_{4'}$	1	–	yes	yes
X_1	1	yes	–	yes
$X_{5'}$	2	–	–	yes
X_3	1	–	yes	–
L_1	1^*	yes	yes	yes
L_3	2^*	yes	yes	yes
L_3	2^*	yes	yes	yes
$L_{2'}$	1^*	–	yes	yes
L_1	1	yes	yes	yes
$L_{2'}$	1	–	–	–
$L_{3'}$	2	–	–	yes
$W_{2'}$	1^*	yes	yes	yes
W_3	2^*	yes	yes	yes
W_1	1^*	yes	yes	yes
$W_{1'}$	1^*	yes	yes	yes
W_3	2	–	yes	yes
$W_{2'}$	1	–	–	yes
W_1	1	yes	–	yes

plus four equivalent sp^3-like hybrids centered near the atom.

C. Silicon

Several authors have previously discussed and computed WFs for silicon and other tetrahedral semiconductors. Some works have focused on the WFs associated with the valence bands,[16,28,43–46] while others have also dealt with the lowest four conduction bands.[29,30]

1. Bond orbitals

A set of eight bond-centered WFs, four bonding and four antibonding, can be obtained by using separate energy windows for each of the two groups, as indicated in Fig. 8(a). Since the valence bands form an isolated group, inside the corresponding window $N_{\mathbf{k}} = N = 4$ throughout the Brillouin zone. Hence there is no freedom for minimizing $\Omega_{\rm I}$, and one can proceed directly with the minimization of $\widetilde{\Omega}$ to compute the "maxloc" WFs, as done in Ref. 16. The resulting bands are essentially indistinguishable from the original ones, since for such a dense k-mesh the interpolation error is very small. The trial orbitals used to start the minimization were bond-centered Gaussians with a root mean-square (r.m.s.) width of 1.89 bohr. The value of the optimal spread was $\Omega = 30.13$ bohr^2, of which 28.39 bohr^2 came from $\Omega_{\rm I}$.

The use of an energy window becomes necessary for the four low-lying empty bands, which are attached to higher bands. As trial orbitals we used an antibonding combination of Gaussians with a r.m.s. width of 1 bohr. Each Gaussian was sitting halfway between one of the two atoms and the center of their common bond. During the minimization $\Omega_{\rm I}$ decreased from 106.76 bohr^2 to 87.47 bohr^2, having reached the minimum in less than 30 steps. (An alternative is to choose the initial subspace at each $\mathbf{k}$ as the lowest four energy eigenstates inside the energy window. This yields an initial $\Omega_{\rm I} = 98.10$ bohr^2, and again the absolute minimum is reached after $\sim$30 steps.) The total spread of the four "maxloc" WFs was $\Omega = 97.49$ bohr^2; as expected,[28] this is considerably larger than for the bonding WFs. Note also that $\widetilde{\Omega}$ accounts for more than 10% of the total spread, whereas for the bonding "maxloc" WFs that number was less than 6%. This is related to the fact that the antibonding WFs are more spread out, causing matrix elements of the type $\langle w_{m\mathbf{R}}|\mathbf{r}|w_{n\mathbf{0}}\rangle$ with $\mathbf{R} \neq \mathbf{0}$ to have larger values. Eq. (15) of Ref. 16 shows that this results in a larger $\widetilde{\Omega}$. The very small contribution of $\widetilde{\Omega}$ to the total spread of the highly localized d-like WFs in copper (less than 1%), as well as the comparatively larger contribution in the interstitial-centered WFs are thus easily understood.

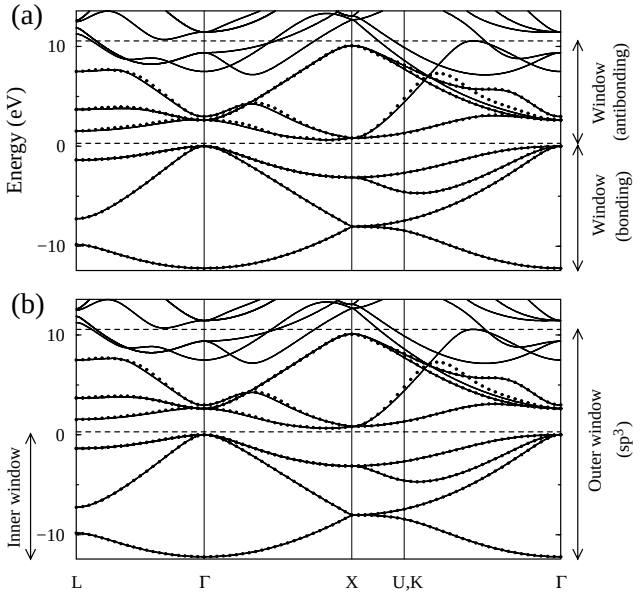

FIG. 8. Solid lines: Original band structure of silicon. Dotted lines: Wannier-interpolated bands. In (a) the valence and low-lying conduction bands are treated separately, which produces four bonding and four antibonding Wannier functions; in (b) they are treated as a single group, which yields eight sp^3-type Wannier functions.

In Fig. 9(a) we present the contour-surface plot of one "maxloc" antibonding WF in silicon. The other three are identical (related to the first by the tetrahedral symmetry operations). Fig. 9(b) shows one of the four identical bonding WFs.

2. sp^3 hybrids

As discussed in Ref. 30, one may instead treat the four valence and four low-lying conduction bands as a single group, which leads to "maxloc" WFs of sp^3 character (Fig. 9(c)). Using our method this may be done as indicated in Fig. 8(b). An outer energy window is chosen which spans the eight bands of interest, and the valence bands are "frozen" inside an inner window; this ensures that they are not affected by the minimization of $\Omega_{\rm I}$, whose only aim is to extract the four low-lying antibonding bands from the conduction band complex. We have started the minimization of $\Omega_{\rm I}$ in two different ways: (i) by projecting eight "atom-centered" sp^3-type combinations of Gaussians, and (ii) by projecting four bond-centered Gaussians plus four antibonding combinations of Gaussians, as done in the previous Section. In both cases the minimization took about 20 steps, taking from 76.04 bohr^2 in the former case and 84.08 bohr^2 in the latter to 63.50 bohr^2. As for the minimization of $\widetilde{\Omega}$, the absolute minimum ($\Omega = 85.41$ bohr^2) was reached only with (i); with (ii) the algorithm became trapped in a local minimum ($\Omega = 101.97$ bohr^2) having the same symmetry as the trial orbitals, with four bonding (anti-

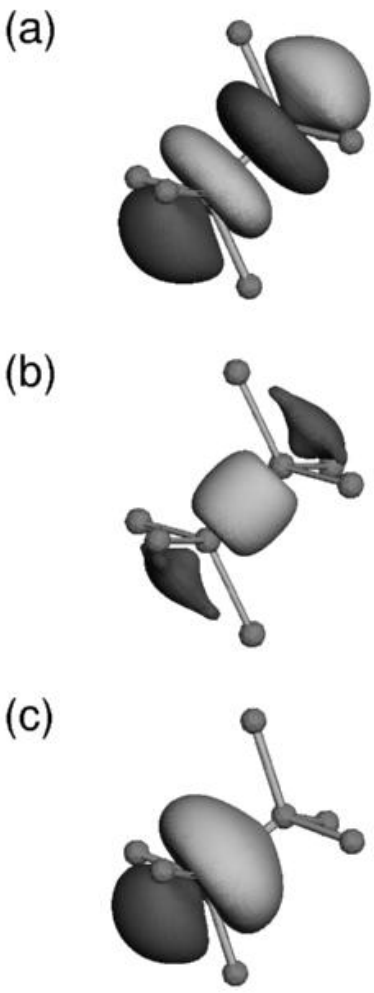

FIG. 9. Contour-surface plots of Wannier functions in silicon. (a) Antibonding, (b) bonding, and (c) sp^3-type. In (a) and (c) the amplitudes are $+0.5/\sqrt{v}$ (light gray) and $-0.5/\sqrt{v}$ (dark gray); in (b) they are $+1.4/\sqrt{v}$ (light gray) and $-0.4/\sqrt{v}$ (dark gray).

bonding) WFs with a spread of 6.37 bohr^2 (19.12 bohr^2) each.

We end this section with the following observation. Suppose we take the four-dimensional valence (bonding) space $\mathcal{S}_4^{(b)}(\mathbf{k})$ together with the optimal four-dimensional antibonding subspace $\mathcal{S}_4^{(a)}(\mathbf{k})$ (Fig. 8(a)) to form an eight-dimensional space $\mathcal{S}_8^{'}(\mathbf{k}) = \mathcal{S}_4^{(b)}(\mathbf{k}) \cup \mathcal{S}_4^{(a)}(\mathbf{k})$. This space has $\Omega_{\rm I} = 63.64$ bohr^2, which is slightly higher than the value 63.50 bohr^2 associated with the optimal subspace $\mathcal{S}_8(\mathbf{k})$ for the eight-band problem with an inner window (Fig. 8(b)). Thus, if we take $\mathcal{S}_8^{'}(\mathbf{k})$ as an initial guess for the minimization of $\Omega_{\rm I}$ in the eight-band problem with an inner window, we will be starting slightly above the absolute minimum. The extra reduction in $\Omega_{\rm I}$ comes about because the functional that is minimized to obtain $\mathcal{S}_8(\mathbf{k})$ contains terms involving overlap between low-lying conduction states at $\mathbf{k}$ and valence states at neighboring $\mathbf{k}+\mathbf{b}$. The wavefunctions relax in response to these extra terms, and consequently the two antibonding subspaces are not exactly the same. However, they are almost identical, and therefore the same is true for the interpolated bands (compare Figs. 8(a) and 8(b)).

V. CONCLUSIONS

We have discussed and implemented a practical method for extracting maximally-localized Wannier functions from entangled energy bands, starting from the Bloch eigenfunctions obtained in a standard electronic structure calculation. Our method is based on a prescription for "disentangling" the bands of interest from the rest of the band complex inside an energy window specified by the user. The idea is to extract a subspace of Bloch-like states whose character varies as little and as smoothly as possible across the Brillouin zone. This is achieved by minimizing a functional which measures the "spillage", or change of character of the subspace across the Brillouin zone. The present scheme can be viewed as an extension of the maximally-localized Wannier function method of Marzari and Vanderbilt,[16] which was designed to deal with isolated groups of bands only. More precisely, it introduces an extra step – the construction of the optimal subspace – which is followed by the determination of the "maxloc" WFs by applying the localization algorithm of Marzari and Vanderbilt to that subspace. The procedure for determining this optimal subspace is both stable and computationally very fast.

Some possible applications of such WFs have been mentioned in the Introduction. Of particular interest is the ability to obtain WFs for the low-lying empty or partially filled bands. For instance, it has been suggested that these could be useful for accurate calculations of the optical properties of semiconducting nanocrystals.[47] Another potential use of the present method could arise in the description of surface states (e.g., Ref. 48), in particular when the surface bands become resonant with the bulk bands. The striking result that we have obtained for the low-lying broad bands of copper, with the WFs being centered at the tetrahedral interstitial sites, suggests that the method may provide insight into the chemistry of transition metal compounds. Also, since the "maxloc" WFs provide a compact interpolation scheme for the band structure, they could be used as part of an efficient algorithm for determining the Fermi surface. Finally, it might be interesting to apply the present ideas to the construction of lattice WFs describing the part of the phonon spectrum relevant for studying structural phase transitions.[49,50]

ACKNOWLEDGMENTS

This work was supported by the NSF Grant DMR-9981193. We would like to thank Dr. Noam Bernstein for providing us with his visualization software *dan*.

[1] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Saunders, New York, 1976).
[2] E. Fradkin, *Field Theories of Condensed Matter Systems* (Addison-Wesley, 1991).
[3] M. S. Hybertsen, M. Schlüter, and N. E. Christensen, Phys. Rev. B **39**, 9028 (1989).
[4] A. K. McMahan, J. F. Annett, and R. M. Martin, Phys. Rev. B **42**, 6268 (1990).
[5] V. I. Anisimov, A. I. Poteryaev, M. A. Korotin, A. O. Anokhin, and G. Kotliar, J. Phys. Condens. Matt. **35**, 7359 (1997).
[6] G. H. Wannier, Phys. Rev. **52**, 191 (1937).
[7] P. Y. Yu and M. Cardona, *Fundamentals of Semiconductors* (Springer, Berlin, 1996).
[8] J. C. Slater, Phys. Rev. **52**, 198 (1937).
[9] S. Goedecker, Rev. Mod. Phys. **71**, 1085 (1999).
[10] R. D. King-Smith and D. Vanderbilt, Phys. Rev. B **47**, 1651 (1993).
[11] R. W. Nunes and D. Vanderbilt, Phys. Rev. Lett. **73**, 712 (1994).
[12] R. Resta and S. Sorella, Phys. Rev. Lett. **82**, 370 (1999).
[13] I. Souza, T. Wilkens, and R. M. Martin, Phys. Rev. B **62**, 1666 (2000).
[14] E. K. Kudinov, Phys. Solid State **41**, 1450 (1999).
[15] E. Koch and S. Goedecker, Solid State Commun. **119**, 105 (2001).
[16] N. Marzari and D. Vanderbilt, Phys. Rev. B **56**, 12 847 (1997).
[17] P. L. Silvestrelli, N. Marzari, D. Vanderbilt, and M. Parrinello, Solid State Communications **107**, 7 (1998); M. Fornari, N. Marzari, M. Peressi, and A. Baldereschi, Comp. Mater. Science **20**, 337 (2001).
[18] N. Marzari and D. Vanderbilt, AIP Conf. Proc. **436**, 146 (1998).
[19] P. L. Silvestrelli and M. Parrinello, Phys. Rev. Lett. **82**, 3308 (1999).
[20] I. Souza, R. M. Martin, N. Marzari, X. Zhao, and D. Vanderbilt, Phys. Rev. B **62**, 15505 (2000).
[21] M. Posternak, A. Baldereschi, S. Massidda, and N. Marzari (in preparation).
[22] O. K. Andersen, A. I. Liechtenstein, O. Jepsen, and F. Paulsen, J. Phys. Chem. Solids, **56**, 1573 (1995).
[23] O. K. Andersen and T. Saha-Dasgupta, Phys. Rev. B **62**, R16219 (2000).
[24] D. A. Goodings and R. Harris, Phys. Rev. **178**, 1189 (1969).
[25] P. Modrak and R. Wojnecki, Phys. Rev. B **36**, 5830 (1987).
[26] P. Modrak, Phys. Rev. B **46**, 15716 (1992).
[27] B. Sporkmann and H. Bross, Phys. Rev. B **49**, 10869 (1994).
[28] H. Teichler, Phys. Stat. Sol. (b) **43**, 307 (1971).
[29] B. Sporkmann and H. Bross, J. Phys.: Condens. Matter **9**, 5593 (1997).
[30] W. Kohn, Phys. Rev. B **7**, 4388 (1973).
[31] H. Bross, Z. Phys. **243**, 311 (1971).
[32] D. Sánchez-Portal, E. Artacho, and J. Soler, Solid State Commun. **95**, 685 (1995); J. Phys.: Condens. Matter **8**, 3859 (1996).
[33] For $\alpha \neq 1$ Eqs. 12 and 18 are not equivalent, except at self-consistency.
[34] P. O. Löwdin, J. Chem. Phys. **18**, 365 (1950).
[35] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in Fortran* (Cambridge, 1988).
[36] J. C. Slater and G. F. Koster, Phys. Rev. **94**, 1498 (1954).
[37] For a single band the interpolated band structure is gauge-invariant, so that the smoothness is independent of the degree of localization of the WF.
[38] N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991).
[39] For the smaller energy windows in Table I the five d-like WFs start to lose symmetry, with their spread no longer being precisely split into two groups.
[40] G. A. Burdick, Phys. Rev. **129**, 138 (1963).
[41] W. A. Harrison, *Electronic Structure and the Properties of Solids* (Freeman, New York, 1980).
[42] D. A. Papaconstantopoulos, *Handbook of the Band Structure of Elemental Solids* (Plenum, New York, 1986).
[43] E. O. Kane and A. B. Kane, Phys. Rev. B **17**, 2691 (1978).
[44] C. Tejedor and J. A. Vergés, Phys. Rev. B **19**, 2283 (1979).
[45] S. Satpathy and Z. Pawlowska, Phys. Stat. Sol. (b) **145**, 555 (1988).
[46] P. Fernández, A. Dal Corso, A. Baldereschi, and F. Mauri, Phys. Rev. B **55**, R1909 (1997).
[47] A. Mizel and M. L. Cohen, Phys. Rev. B **56**, 6737 (1997); Solid State Commun. **104**, 401 (1997).
[48] G. Santoro, S. Scandolo, and E. Tosatti, Phys. Rev. B **59**, 1891 (1999); C. S. Hellberg and S. C. Erwin, Phys. Rev. Lett. **83**, 1003 (1999).
[49] K. M. Rabe and U. V. Waghmare, Phys. Rev. B **52**, 13236 (1995).
[50] J. Íñiguez, A. García, and J. M. Pérez-Mato, Phys. Rev. B **61**, 3127 (2000).