Microscopic electronic wave function and interactions between quasiparticles in empirical tight-binding theory

R. Benchamekh,[1,2] F. Raouafi,[2] J. Even,[3] F. Ben Cheikh Larbi,[2] P. Voisin,[1] and J.-M. Jancu[3]

[1]*CNRS-Laboratoire de Photonique et de Nanostructures, Route de Nozay, F-91460 Marcoussis, France*

[2]*Laboratoire de Physico-chimie des Microstructures et Micro-systèmes, Institut Préparatoire aux Etudes Scientifiques et Techniques, BP51, 2070 La Marsa, Tunisia*

[3]*FOTON, Université Européenne de Bretagne, INSA-Rennes and CNRS, Rennes, France*

(Received 28 March 2013; revised manuscript received 30 December 2014; published 15 January 2015)

A procedure to obtain single-electron wave functions within the tight-binding formalism is proposed. It is based on linear combinations of Slater-type orbitals whose screening coefficients are extracted from the optical matrix elements of the tight-binding Hamiltonian. Bloch functions obtained for zinc-blende semiconductors in the extended-basis $spds^*$ tight-binding model demonstrate very good agreement with first-principles wave functions. We apply this method to the calculation of the electron-hole exchange interaction, and obtain the dispersion of excitonic fine structure in bulk GaAs. Beyond semiconductor nanostructures, this work is a fundamental step toward modeling many-body effects from post-processing single-particle wave functions within the tight-binding theory.

DOI: 10.1103/PhysRevB.91.045118 PACS number(s): 71.15.−m, 71.35.−y, 71.70.Gm

I. INTRODUCTION

Tight-binding is widely used as a conceptual frame to account for the kinetic energy operator in solid state theory. It served as a basis for major contributions such as Anderson's strong localization [1] and magnetic impurity [2] theories, Hubbard's model of interacting electrons [3], and many others [4–6]. In these theories, interactions are generally introduced using symmetry considerations and *ad hoc* parameters, and the goal is to have the simplest model integrating the physically relevant features of hopping and interaction integrals while getting rid of the complexities of underlying atomic physics. On the other hand, empirical-parameter tight-binding (EPTB) is known as a powerful modeling technique for the electronic structure of semiconductors, metals, and all kinds of nanoscale structures and devices. The systematic procedure for constructing the EPTB Hamiltonian was discussed in a seminal paper by Slater and Koster (SK) in 1954 [7], but it is only many years later that computers have allowed systematic implementations [8,9]. A major step was achieved in the late 1990s with the development of EPTB models using an extended $spds^*$ orbital basis [10,11] and allowing accurate full-band representation of single-particle states. However, when it comes to calculating short-range interactions between quasiparticles, EPTB models (like the $\mathbf{k}\cdot\mathbf{p}$ theory itself) are hampered by the lack of an explicit orbital basis. This is a strong limitation for the use of advanced EPTB schemes to explore highly topical problems of strongly correlated electron systems. In this paper, we present a method that reconciles the "conceptual frame" and "modeling tool" faces of tight-binding, by self-consistently determining the orbital basis out of the EPTB Hamiltonian. We thus obtain the local wave functions, which allows parameter-free calculation of short-range interactions. We illustrate this by calculating electron-hole exchange and the fine structure of excitons in GaAs. Beyond semiconductors, this method can be used for many different materials and can handle million-atom supercells that are still out of reach of first-principle methods.

II. SELF-CONSISTENT PROJECTION BASIS FOR THE EPTB HAMILTONIAN

In the SK formalism, the crystal potential is approximated as a sum of spherically symmetrical potentials around each atom. This allows the electronic wave functions $\Psi_{\alpha\mathbf{k}}$, where α stands for the band indices, to be developed on a set of Bloch sums $\Phi_{lm\mathbf{k}}$ of atomic-like orbitals (called the Löwdin orbitals) ψ_{lmj}, where ψ_{lmj} is the mth orbital on the lth atom in the jth unit cell:

$$\Psi_{\alpha\mathbf{k}} = \sum_{m,l} C^{\alpha}_{ml}\Phi_{lm\mathbf{k}}, \tag{1}$$

$$\Phi_{lm\mathbf{k}} = \frac{1}{\sqrt{N}}\sum_{j}\exp(i\mathbf{r}_{jl}\cdot\mathbf{k})\psi_{ml}(\mathbf{r}-\mathbf{r}_{jl}) \tag{2}$$

The Löwdin orbitals have well defined angular properties, but unknown radial dependencies. The Hamiltonian matrix elements between them are treated as "disposable constants" with which one can fit band structures that have been experimentally determined or calculated using more accurate techniques. The fit is performed in k-space, removing any necessity to further characterize the local wave functions in real space. Very interestingly, following a method introduced by Boykin and Vogl, interaction with the electromagnetic field can be built-in using a derivation of the Hamiltonian matrix elements in momentum space [12]:

$$p_{lm,l'm'} = i\hbar\langle\Phi_{lm}\mid\nabla_k H\mid\Phi_{l'm'}\rangle. \tag{3}$$

Optical properties can be consequently calculated in EPTB models without adding parameters. Although this method misses intra-atomic matrix elements, good accuracy is obtained provided that the orbital basis is rich enough [13]. In this context, the $spds^*$ TB model is known to give a description of dielectric properties equivalent to best *ab initio* calculations within the one-electron approximation [14]. Altogether, the model's major qualities are the transferability of parameters

from bulk materials to nanostructures, the unique ability to describe with the same accuracy electronic properties in any region of the Brillouin zone, and the capacity to handle large supercells. However, interactions (in particular, short-range interactions) between quasiparticles involve the local wave functions, and cannot be calculated in this frame. Existing calculations of Coulomb matrix elements use approximations on the radial dependence of the basis orbitals [15,16].

In order to solve this theoretical issue that has remained open ever since the seminal work of Slater and Koster, we start with a trial set $\mathcal{B}$ of $spds^*$ basis functions in the form of normalized Slater-type orbitals (STOs) $\phi_\mu(r) = \sqrt{(2\alpha)^{2n+1}/(2n)!}\, Y_\mu(\theta,\phi) r^{n-1} e^{-\alpha r}$, where n is the first quantum number, α is a screening parameter [17], and μ stands for the symmetry of the orbital. STOs are largely employed in quantum chemistry but do not fulfill the orthogonality condition, since finite overlap exists between two STOs localized at different sites of the crystal. Following the notation of Ref. [18], we define the overlap integral between two STOs of symmetry μ and ν and localized at the lattice points $\mathbf{R}_1$ and $\mathbf{R}_2$ respectively as

$$S_{\mu\mathbf{R}_1,\nu\mathbf{R}_2} = \int \phi_\mu(\mathbf{r}-\mathbf{R}_1)\phi_\nu(\mathbf{r}-\mathbf{R}_2)d\tau - \delta_{\mu\nu}\delta_{\mathbf{R}_1\mathbf{R}_2}, \quad (4)$$

where $d\tau$ is the volume element. Because of the periodicity of the crystal, the overlap integral depends only on the difference between $\mathbf{R}_2$ and $\mathbf{R}_1$ and thus can be written as $S_{\mu,\nu}(\mathbf{R}_2 - \mathbf{R}_1)$.

First, the orbital overlap matrix $\mathbf{S}$ is calculated including all orbitals up to a cutoff distance R_0 that must be taken large enough so that overlap with remote atoms be negligible. Thanks to the nice properties of STOs, this step is done analytically, and in practice we found that, for $\alpha = 0.5$ Å^{-1}, overlap with neighbors located farther than 3 lattice parameters (17 Å) can safely be neglected. Note that, unlike real atomic orbitals, s and s^* on-site STOs are not orthogonal. This small difficulty is easily solved by substituting s^* with a Gramm-Schmitt combination $\tilde{s}^* = (s^* - \langle s|s^*\rangle s)/\sqrt{1-\langle s|s^*\rangle^2}$. Then, the functions ψ_μ of the orthogonal basis $\mathcal{B}_{orth}$ can be expressed in terms of the functions ϕ_μ of the nonorthogonal basis $\mathcal{B}$ using the Löwdin orthogonalization procedure [18]:

$$\{\psi\} = \{\phi\}\cdot(\mathbb{1}+\mathbf{S})^{-\frac{1}{2}} \quad (5)$$

or

$$\begin{aligned}\psi_\lambda(\mathbf{r}-\mathbf{R}_1) &= \phi_\lambda(\mathbf{r}-\mathbf{R}_1) - \frac{1}{2}\sum_{\mu,\mathbf{R}_2}\phi_\mu(\mathbf{r}-\mathbf{R}_2)S_{\mu\mathbf{R}_2,\lambda\mathbf{R}_1} \\ &+ \frac{3}{8}\sum_{\mu,\mathbf{R}_2}\sum_{\nu,\mathbf{R}_3}\phi_\mu(\mathbf{r}-\mathbf{R}_2)S_{\mu\mathbf{R}_2,\nu\mathbf{R}_3}S_{\nu\mathbf{R}_3,\lambda\mathbf{R}_1} - + \cdots, \end{aligned} \quad (6)$$

where $\{\psi\}$ and $\{\phi\}$ denote the row matrices formed by the elements ψ_μ and ϕ_μ, respectively. The orthogonalized STOs will serve as trial functions for the unknown Löwdin orbitals. The expansion of electronic eigenfunctions (Bloch functions) in the basis $\mathcal{B}$ is obtained by multiplying the eigenvectors of the $sp^3d^5s^*$ Hamiltonian matrix by the matrix $(\mathbb{1}+\mathbf{S})^{\frac{1}{2}}$, which provides definitely their representation in real space. Then the momentum matrix elements are calculated *in real space* from the Bloch sums by the relation

$$p_{lm,l'm'} = i\hbar\langle\Phi_{lm} \mid \nabla_r \mid \Phi_{l'm'}\rangle \quad (7)$$

This derivation involves a sum of matrix elements between two STOs that are calculated analytically [19]. Finally the screening parameters are fitted into a genetic algorithm until the optical matrix elements calculated in real space compare satisfactorily with those derived in k-space from the electronic Hamiltonian. In the end, the optical matrix element between two electronic bands denoted by α and β is obtained by

$$P_{\alpha,\beta} = \sum_{l,m}\sum_{l'm'} C^\alpha_{lm}\, p_{lm,l'm'}\, C^\beta_{l'm'}. \quad (8)$$

We note that one can derive similar approaches by considering other atomic like orbitals (e.g., Gaussians instead of the radial part of STOs) or by using another orthogonalization procedure. However, besides the mathematical properties of STOs allowing analytical calculation of almost every observable, STOs are close to the envelopes of real atomic orbitals and the latter can be further approached by considering linear combination of STOs. On the other hand, the choice of Löwdin's method rather than any other orthogonalization procedure is clearly justified since, as proven by Slater and Koster in the appendix of Ref. [7], Löwdin's method has a unique merit that an orthogonal basis set remains as close as possible to the original nonorthogonal set and retains their symmetry characteristics.

III. PRACTICAL IMPLEMENTATION AND DISCUSSION

In practice however, the 9150×9150 $\mathbf{S}$ matrix obtained for $R_0 = 17$ Å contains considerably redundant information (for instance, overlap of two given orbitals on neighboring sites appears 3528 times) and its large size makes its inversion computationally difficult. Fortunately, one can take advantage of the crystal periodicity and reduce the size of the matrix to be inverted. For this, we consider Bloch sums of STOs:

$$\kappa_{\mu,\mathbf{k}} = \frac{1}{\sqrt{N}}\sum_{\mathbf{R}} e^{i\mathbf{R}\cdot\mathbf{k}}\phi_\mu(\mathbf{r}-\mathbf{R}). \quad (9)$$

In the $spds^*$ model, the $\{\kappa_\mathbf{k}\}$ set contains 40 elements for each $\mathbf{k}$. We construct the 40×40 overlap matrix $\tilde{\mathbf{S}}$ between these Bloch sums of STOs:

$$\tilde{S}_{\mu\nu} = \sum_{\mathbf{R}} e^{i\mathbf{R}\cdot\mathbf{k}} S_{\mu\nu}(\mathbf{R}). \quad (10)$$

Then the orthogonal Bloch sums $\chi_{\mu,\mathbf{k}}$ are obtained by Löwdin's orthogonalization:

$$\{\chi_\mathbf{k}\} = \{\kappa_\mathbf{k}\}\cdot(\mathbb{1}+\tilde{\mathbf{S}})^{-\frac{1}{2}}, \quad (11)$$

where $\{\chi_\mathbf{k}\}$ and $\{\kappa_\mathbf{k}\}$ denote the row matrices formed by the elements $\chi_{\mu,\mathbf{k}}$ and $\kappa_{\mu,\mathbf{k}}$, respectively. Although this may be not obvious at first glance, the orthogonalized Bloch sums of STOs are identical to Bloch sums of orthogonalized STOs. Indeed, multiplying Eq. (6) by $\frac{1}{\sqrt{N}}e^{i\mathbf{R}_1\cdot\mathbf{k}}$, summing over lattice points

TABLE I. Optimized Slater orbital screening coefficients ($Å^{-1}$) for gallium (Ga) and arsenic (As), compared with Slater's atomic screening coefficients [17].

	Ga		As	
Orbital	Ref. [17]	This work	Ref. [17]	This work
$4s$	1.35	1.83	1.7	1.94
$4p$	1.35	1.77	1.7	1.79
$4d$	0.27	0.93	0.27	0.96
$5s$	0.32	1.64	0.4	1.74

$\mathbf{R}_1$, and using the fact that the overlap integrals $S_{\mu,\nu}(\mathbf{R}_1,\mathbf{R}_2)$ depend only on $\mathbf{R}_2 - \mathbf{R}_1$, we obtain

$$\frac{1}{\sqrt{N}}\sum_{\mathbf{R}_1} e^{i\mathbf{R}_1\cdot\mathbf{k}}\psi_\lambda(\mathbf{r}-\mathbf{R}_1) = \kappa_{\lambda,\mathbf{k}} - \frac{1}{2}\sum_\mu \kappa_{\mu,\mathbf{k}}\tilde{S}_{\mu,\lambda} + \frac{3}{8}\sum_\mu\sum_\nu \kappa_{\mu,\mathbf{k}}\tilde{S}_{\mu,\nu}\tilde{S}_{\nu,\lambda} - + \cdots \quad (12)$$

and thus

$$\frac{1}{\sqrt{N}}\sum_{\mathbf{R}_1} e^{i\mathbf{R}_1\cdot\mathbf{k}}\psi_\lambda(\mathbf{r}-\mathbf{R}_1) = \chi_{\lambda,\mathbf{k}}. \quad (13)$$

Since the procedure sketched in Sec. II uses the Bloch sums rather than the individual on-site orbitals, it can therefore be implemented at much lower computational cost using the $\tilde{\mathbf{S}}$ instead of the $\mathbf{S}$ matrix. Once the screening parameters have been optimized though the genetic algorithm (see Table I), we can invert the full $\mathbf{S}$ matrix and get the Löwdin orbitals. The effect of orthogonalization on the different basis functions is illustrated with the case of arsenic in GaAs in Figs. 1, 2, and 3. The final set of screening parameters is used in the calculation. It can be seen that the STOs are not too severely modified by the orthogonalization procedure, which indicates that they are a fair zeroth-order guess.

As for the determination of interband matrix elements, the consideration of different bands and different high symmetry points in the Brillouin zone provides more than necessary information for the fit convergence from a computational point of view. However, we observe that somewhat different sets of screening parameters can give similar "fitness" parameters. This problem is linked with the fact that upper-band dispersions are necessarily incorrect, due to the noncompleteness of finite basis. Hence, optical matrix elements involving s^* and d orbitals are poorly represented in k-space, and using them to fit screening parameters can lead to an unphysical optimum in the algorithm. Noncompleteness is a fundamental difficulty of the EPTB method. On the other hand, the $d(\Gamma_{12})$ eigenstate (notation refers to the specific case of zinc-blende crystals) is known to be a nearly-free electron state weakly perturbed by the crystal potential, and its wave function can be accurately determined from empirical pseudopotential or *ab initio* calculations. Constraining the parameter space in such a way that the $d(\Gamma_{12})$ orbital agrees with independent empirical

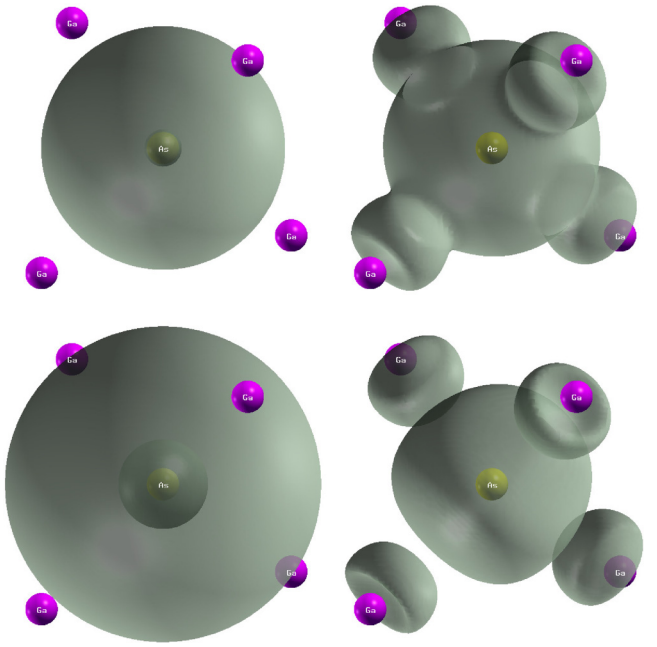

FIG. 1. (Color online) Arsenic s (top) and s^* (bottom) Slater orbitals before (left) and after (right) orthogonalization.

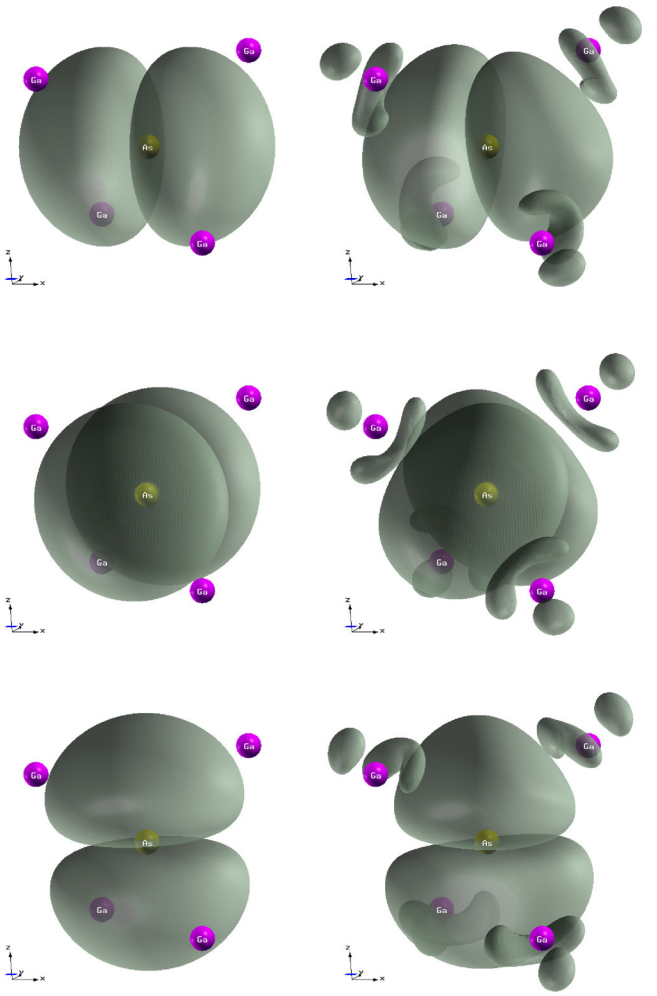

FIG. 2. (Color online) Arsenic p_x, p_y, and p_z Slater orbitals before (left) and after (right) orthogonalization.

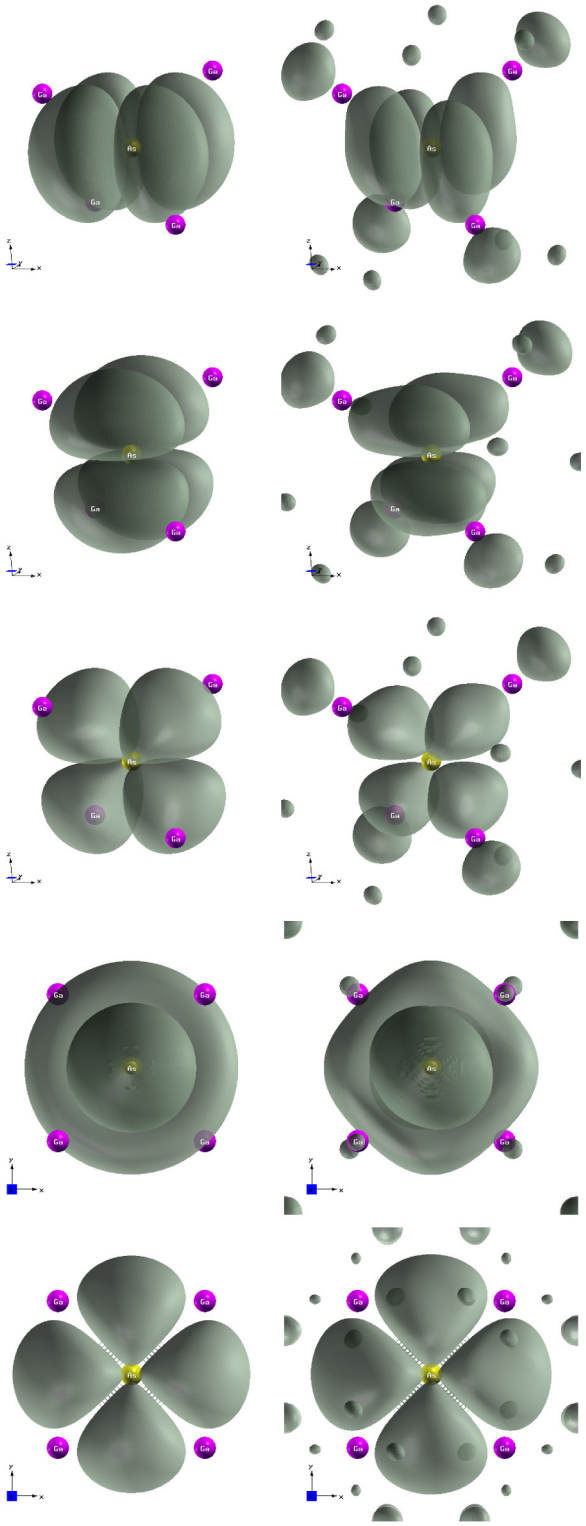

FIG. 3. (Color online) Arsenic d_{xy}, d_{yz}, d_{xz}, $d_{3z^2-r^2}$, and $d_{x^2-y^2}$ Slater orbitals before (left) and after (right) orthogonalization.

pseudo-potential or *ab initio* calculations cures the problem of underdetermination of the set of screening parameters. In Table I, we compare the optimized screening coefficients for GaAs with Slater's atomic screening coefficients [17]. Due to the self-consistent procedure, they obviously depend on the Hamiltonian parameters, which are given in Tables II and III

TABLE II. Tight-binding parameters for Ge (eV).

a	5.6500	E_s	−3.2967
E_{s^*}	19.1725	E_p	4.6560
E_d	13.0143	$ss\sigma$	−1.5002
$ss^*\sigma$	−1.9206	$s^*s^*\sigma$	−3.6029
$sp\sigma$	2.7985	$s^*p\sigma$	2.8176
$sd\sigma$	−2.8028	$s^*d\sigma$	−0.6209
$pp\sigma$	4.2540	$pp\pi$	−1.6510
$pd\sigma$	−2.2138	$pd\pi$	1.9001
$dd\sigma$	−1.2171	$dd\pi$	2.5054
$dd\delta$	−2.1389	$\Delta/3$	0.12742

for Ge and GaAs respectively. These parameters are a slightly reworked version of the classical set of Ref. [10].

We applied the procedure explained above to the prototype systems of Ge and GaAs. The electron configuration of Ge is $[\mathrm{Ar}]3d^{10}4s^24p^2$. In the $spds^*$ model, the deep $3d$ states are discarded and the basis is formed by the orbitals $4s$, the three orbitals $4p$, the five empty orbitals $4d$ and the empty orbitals $5s$. When building the STO basis $\mathcal{B}$, we keep fixed the first quantum number n of these orbitals and introduce one adjustable screening parameter α for each symmetry type. Alternatively, as often done in quantum chemistry, we can improve parametrical flexibility by considering that each element of the starting basis is a linear combination of q STOs instead of one. This does not change much the model, but increases to $4q$ the number of fitted parameters. For GaAs, since there are two different atoms in the unit cell, the number of parameters is twice that for Ge. The fitted screening parameters for Ga and As are given in Table I, and contrasted with the Slater atomic screening constants. Those obtained independently for Ge are close to averaged values for Ga and As. At the end of the fitting procedure, the global discrepancy on the sum of all interband matrix elements, calculated at the Γ, X, and L points of the Brillouin zone, is less than 15% with one Slater orbital per atomic state and less than 7% with a linear combination of two Slater orbitals for each atomic state. By changing the relative weights of different spectral or Brillouin zone regions

TABLE III. Tight-binding parameters for GaAs (eV).

a	5.6500	E_s^a	−5.9820
$E_{s^*}^a$	19.4477	E_s^c	−0.3803
$E_{s^*}^c$	19.4548	E_p^a	3.3087
E_d^a	13.2015	E_p^c	6.3801
E_d^c	13.2055	$ss\sigma$	−1.6874
$s_as_c^*\sigma$	−1.5212	$s_a^*s_c\sigma$	−2.1058
$s^*s^*\sigma$	−3.7170	$s_ap_c\sigma$	2.8845
$s_cp_a\sigma$	2.8902	$s_a^*p_c\sigma$	2.5294
$s_c^*p_a\sigma$	2.3883	$s_ad_c\sigma$	−2.8716
$s_cd_a\sigma$	−2.2801	$s_a^*d_c\sigma$	−0.6568
$s_c^*d_a\sigma$	−0.6113	$pp\sigma$	4.4047
$pp\pi$	−1.4470	$p_ad_c\sigma$	−1.6034
$p_cd_a\sigma$	−1.6260	$p_ad_c\pi$	1.8422
$p_cd_a\pi$	2.1420	$dd\sigma$	−1.0884
$dd\pi$	2.1560	$dd\delta$	−1.8607
$\Delta_a/3$	0.1745	$\Delta_c/3$	0.0408

TABLE IV. Main interband matrix elements (in eV Å) for Ge and GaAs at the Γ point, calculated from differents models.

		WF1[a]	WF2[a]	LDA[b]	Hamiltonian[c]
Ge	P_0	7.69	10.18	8.49	10.14
	Q_0	8.29	8.42	7.32	8.70
GaAs	P_0	7.38	9.88	8.35	9.82
	P_1	0.80	0.93	1.38	0.11
	Q_0	8.16	8.31	7.37	8.72

[a]Real space calculation from TB wave functions with respectively one and two Slater orbital for each basis element.
[b]Real space calculation from LDA wave functions, ABINIT code.
[c]Calculated from Hamiltonian derivation.

in the genetic algorithm fitness function, the discrepancy at, e.g., Γ can be minimized down to the percent range. The residual discrepancy has three distinct physical origins: (i) the difference between orthogonalized STOs and the actual Löwdin orbitals, (ii) the lack of completeness of the $spds^*$ basis, and (iii) the missing intra-atomic contribution in the k-space derivation method. Table IV shows the main interband momentum elements $P_0 \equiv -i\langle s_c|p_x|x_v\rangle$, $P_1 \equiv -i\langle s_c|p_x|x_c\rangle$ and $Q_0 \equiv -i\langle x_c|p_y|z_v\rangle$ obtained for Ge and GaAs at the Γ point. Agreement is very good, but a discrepancy observed for the weak matrix element P_1, for which real space calculations agree, but differ from the Hamiltonian derivation value [20]. This might be a trace of the methodological limitation relative to intra-atomic contributions. While further work is required to explore the method limitations and improve the results, the present achievement is already sufficient for most practical purposes. Note that the interband momentum elements for GaAs calculated using the original Slater screening parameters of Ref. [17] are $P_0 = 1.50$ eV Å and $Q_0 = 5.34$ eV Å. These values are far from experimental ones and prove the need to fit the screening parameters.

Once screening parameters best reproducing the interband matrix elements are obtained, the different Bloch functions can be plotted and compared with *ab initio* calculations. The latter were performed using the ABINIT code [21,22] in the local density approximation (LDA), completed by self-consistent GW correction. Figures 4 and 5 show (110) plane isodensity contours of wave functions in bulk GaAs at the Brillouin zone center. Figure 4 shows valence band states s_v and $y'_v \equiv y_v - x_v$, in both TB and *ab initio* calculations. The overall quantitative agreement is very good, since the overlap between TB and *ab initio* densities is always better than 95%. Yet, TB wave functions appear somewhat less localized in the sense that they have larger density in regions where the *ab initio* density is almost zero, and the Ga/As asymmetry is more pronounced in the ABINIT result. The most significant difference is for the deep s_v state near the atomic sites, for which TB density is significantly smaller. This probably reflects the difference of the projection basis between the two models: TB wave functions are expanded in the basis of Slater orbitals which have a node at the atomic sites while the ABINIT wave functions are expanded in a basis of plane-wave functions that may be maximum on the atomic sites. The wave-vector cut-off used in the ABINIT calculations is important, because this approach

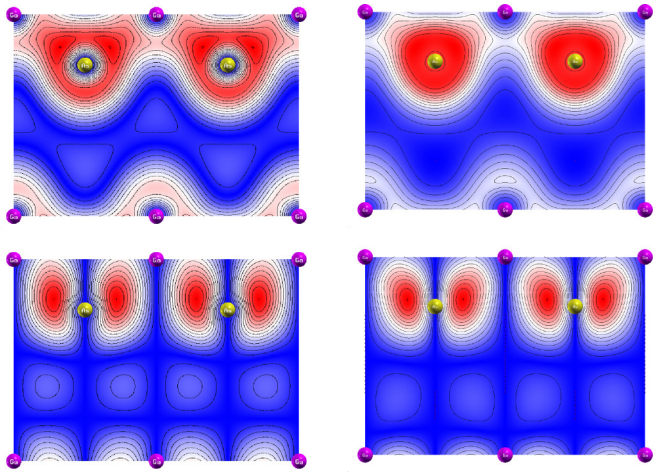

FIG. 4. (Color online) Isodensity contours of the S and $Y' = X + Y$ valence Bloch function at the zone center in bulk GaAs in the plane (110). TB calculation (left) is compared with ABINIT calculations (right).

cannot describe the region located less than $1/k_{\text{cutoff}}$ from atomic sites. Yet, the cutoff does not suffice to explain the observed difference. In our TB approach, a finite on-site value for the s_v state results from the contribution of neighboring atoms and is quite sensitive to the STO screening parameters.

In Fig. 5, we show the conduction Bloch functions also calculated with the same two models. Again TB wave functions are very similar to those calculated in the LDA+GW approximation, except for a significant difference for the s_c

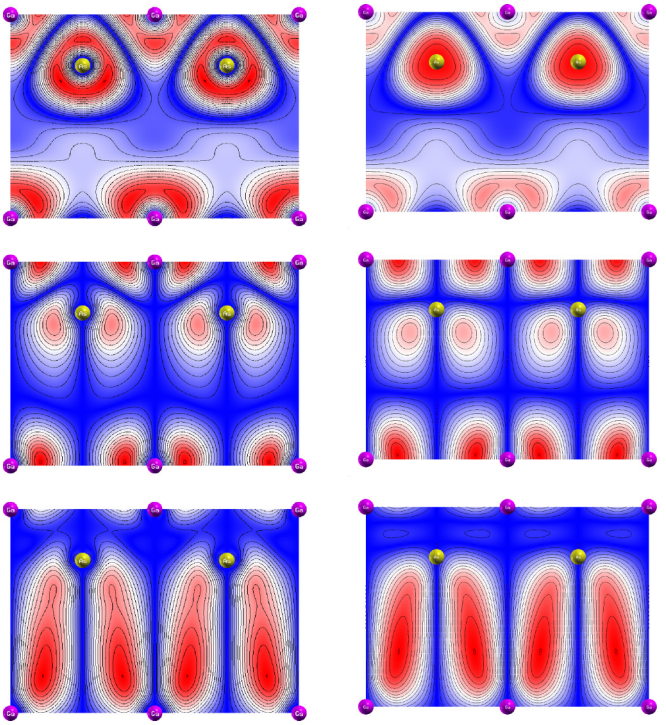

FIG. 5. (Color online) Isodensity contours in the (110) plane for the S, Y', and $D(\Gamma_{12})$ conduction Bloch functions at the zone center in bulk GaAs TB calculation (left) and ABINIT calculations (right).

state density in the vicinity of atomic sites. In order to clarify this issue, we used the SIESTA code, which is based on DFT expanded in a strictly localized orbitals set. SIESTA results for s_c actually agree very well with our TB results (Appendix A).

We note that electron hyperfine interaction constants, that are well documented, scale as ${s_c}^2(r=0)$ and could serve as a quantitative test. A most striking result is the TB ability to reproduce the wave functions of the nearly free electron states s^* and d.

IV. EXCITON FINE STRUCTURE

The major interest of having a real space representation of wave functions is the ability to study many-body problems. In order to illustrate the potential of our wavefunction derivation method, we calculate the binding energy and the fine structure of excitons in bulk GaAs by adding the Coulomb interaction term to the single-particle Hamiltonian. The Coulomb interaction can be specified by the matrix elements

$$\langle \alpha', \mathbf{k}'_e; \beta', \mathbf{k}'_h | U^{eh} | \alpha, \mathbf{k}_e; \beta, \mathbf{k}_h \rangle, \tag{14}$$

where

$$U^{eh} = e^2/\kappa |\mathbf{r}_e - \mathbf{r}_h|.$$

Here, $|\alpha, \mathbf{k}_e; \beta, \mathbf{k}_h\rangle$ is the two-particule excited state, κ is the permittivity, and $\Psi_{\alpha \mathbf{k}_e} \equiv \langle \mathbf{r} | \alpha, \mathbf{k}_e \rangle$ and $\Psi_{\beta \mathbf{k}_h} \equiv \langle \mathbf{r} | \beta, \mathbf{k}_h \rangle$ are the Bloch wave functions in electron and hole representations, respectively [23,24]. The present expansion of Bloch functions as linear combinations of Slater orbitals allows to expand the electron-hole interaction in terms of Coulomb matrix elements between STOs:

$$\begin{aligned} &V_{l_1 m_1, l_2 m_2, l_3 m_3, l_4 m_4} \\ &\quad = \langle \phi_{l_1 m_1}(\mathbf{r} - \mathbf{R}_1), \phi_{l_2 m_2}(\mathbf{r} - \mathbf{R}_2) | U^{eh} | \phi_{l_3 m_3}(\mathbf{r} - \mathbf{R}_3), \\ &\qquad \times \phi_{l_4 m_4}(\mathbf{r} - \mathbf{R}_4) \rangle. \end{aligned} \tag{15}$$

Restricting the expansion to two-center contributions ($\mathbf{R}_1 = \mathbf{R}_3$ and $\mathbf{R}_2 = \mathbf{R}_4$), the evaluation of the integrals can be done quasi-analytically using the expansion of the Coulomb potential in terms of spherical harmonics centered on the same site when $\mathbf{R}_1 = \mathbf{R}_2$ [25], and a bipolar expansion when $\mathbf{R}_1 \neq \mathbf{R}_2$ [26]. Following the approach used in Refs. [27,28], we introduce a r-dependent dielectric constant such that short-range (on-site) interaction is unscreened, while long range interaction is subject to standard dielectric screening. Then, we solve the Bethe-Salpeter equation (BSE), expressed in terms electron and hole single-particle energies and electron-hole interactions calculated using the tight-binding wave functions. The BSE is an eigenvalue problem of infinite dimensionality:

$$\begin{aligned} &(E_{c,\mathbf{k}+\mathbf{Q}/2} - E_{v,\mathbf{k}-\mathbf{Q}/2}) A_{vc\mathbf{k}} \\ &\quad + \int_{V_{BZ}} d^3k' \sum_{v',c'} \langle vc\mathbf{k} | U^{eh} | v'c'\mathbf{k}' \rangle A_{v'c'\mathbf{k}'} = \Omega_S A_{vc\mathbf{k}}, \end{aligned}$$

where $E_{c,\mathbf{k}+\mathbf{Q}/2}$ and $E_{v,\mathbf{k}-\mathbf{Q}/2}$ are the electron and hole energies respectively. Resolution of BSE gives the exciton wavefunction components $A_{vc\mathbf{k}}$ and the excitation energies Ω_S. To make the problem tractable, continuous integration with respect to $\mathbf{k}'$ was replaced by a discrete scheme. Following the procedure described in Refs. [29,30], to calculate the exciton spectra and binding energy at the Γ point, the integration was performed over a small region near the position of the band extrema ($|\mathbf{k}| < 0.015$ a.u.). This region was divided into a $11\times11\times11$ uniform grid. For an exciton wave vector $\mathbf{Q} = 0$, we find an excitonic binding energy $E_b = 4.75$ meV. In addition, the eightfold degenerate $\Gamma_{8v}\times\Gamma_{6c}$ fundamental excitonic transition is split by short range exchange interaction into one twofold, and two threefold degenerate excitons. The twofold and threefold $J = 2$ "dark excitons" are split by $\delta_{\text{anis}} = 0.02$ μeV. This anisotropy splitting is due to the zinc-blende structure which does not allow more than threefold degeneracy. We expectedly find a very small value for δ_{anis}. The $J = 1$ "bright exciton" threefold state is separated from the $J = 2$ states by the short range exchange splitting Δ_{exc}. We get $\Delta_{\text{exc}} = 20.6$ μeV, in agreement with recent experimental

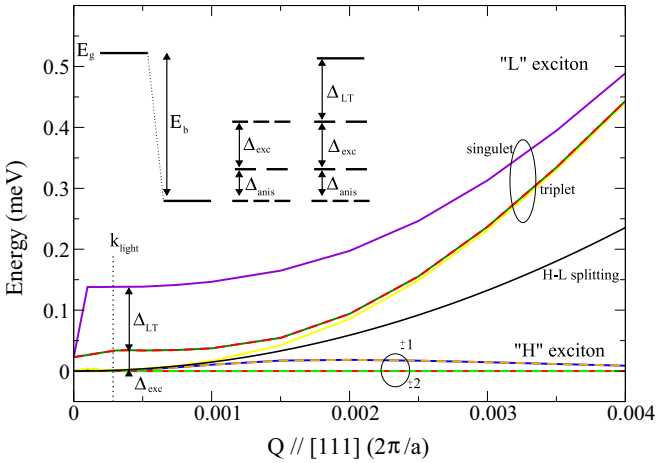

FIG. 6. (Color online) Dispersion of the exciton states of bulk GaAs for $\mathbf{Q}$ along the [111] direction. The long range exchange is exactly zero at $\mathbf{Q} = 0$, and builds up for very small values of $\mathbf{Q}$, for which convergency is more difficult to ensure. The wave vector of light, for which strong polariton features would add to the present picture, is indicated with a vertical line. The inset shows schematically the different contributions to the fine structure of $\Gamma_{8v}\times\Gamma_{6c}$ fundamental exciton.

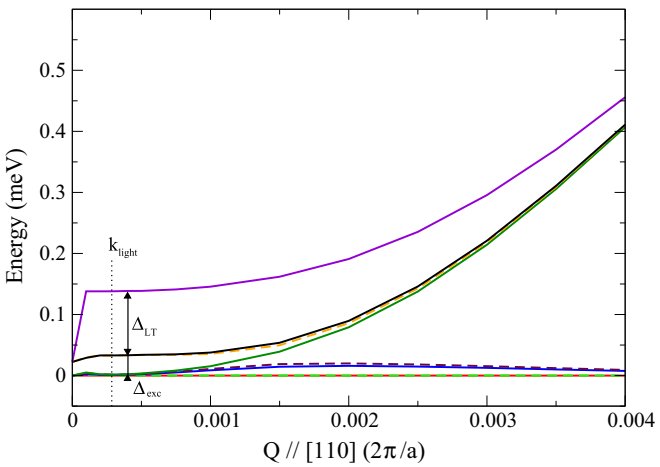

FIG. 7. (Color online) Dispersion of the exciton states of bulk GaAs for $\mathbf{Q}$ along the [110] direction.

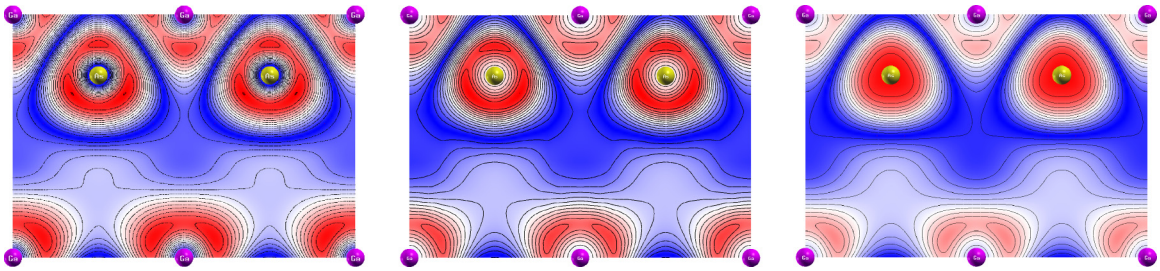

FIG. 8. (Color online) Isodensity contours in the (110) plane for the s_c conduction Bloch function at the zone center in bulk GaAs. TB calculation (left), SIESTA calculations (center), and ABINIT calculations (right).

determination [31]. When one moves away slightly from $\mathbf{Q} = 0$, the $J = 1$ excitons are further split by the long range exchange interaction into twofold degenerate, optically active transverse excitons, and a longitudinal exciton. The energy difference corresponds to the longitudinal-transverse splitting Δ_{LT}, for which we find a value $\Delta_{LT} = 105.3\,\mu$eV in very good agreement with the well documented experimental value.

Then, we examine the evolution of the exciton fine structure as a function of the exciton wave vector $\mathbf{Q}$. Figure 6 shows the calculated dispersion curves. For large $\mathbf{Q}$, when the heavy-hole light-hole splitting becomes larger than exciton binding energy, the exciton splits into a "heavy" exciton formed of two twofold degenerate states and a "light" exciton formed of one threefold degenerate state and one singlet state. Our calculation shows how energy levels interpolate between the small and large $\mathbf{Q}$ regimes. Finally, we note that when $\mathbf{Q}$ is along the [110] direction (Fig. 7), our results show the full details of exciton state spin splittings, including both contributions of electron and hole spin splittings. Note, however, that terms linear in k in hole dispersion are not present in the current $spds^*$ model.

V. CONCLUSION

In conclusion, we have devised a method that allows self-consistent definition of local wavefunctions within the EPTB theory, and successfully used bulk exciton fine-structure as a parameter-free testbed. Extension to nanostructures is straightforward as long as bulk screening parameters are, like other tight binding parameters, transferable to nanostructures. While some fundamental aspects of the method, like the problem of on-site optical matrix elements, still require further clarification, this approach opens a route towards reconciling tight-binding and predictive evaluation of interactions between quasi-particles.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Mikhail Nestoklon and Prof. Pawel Hawrylak for fruitful discussions.

APPENDIX A: SIESTA VERSUS ABINIT AND TB WAVEFUNCTIONS

As discussed in the article, the main differences between the tight-binding wave functions and those obtained using the ABINIT code were observed for s-symmetry states in the

TABLE V. Tight-binding parameters for Si (eV).

a	5.4300	E_s	−2.0386
E_{s^*}	19.9699	E_p	5.0669
E_d	14.8323	$ss\sigma$	−1.8885
$ss^*\sigma$	−1.5103	$s^*s^*\sigma$	−3.6932
$sp\sigma$	2.9607	$s^*p\sigma$	3.5346
$sd\sigma$	−2.5344	$s^*d\sigma$	−2.0505
$pp\sigma$	4.3649	$pp\pi$	−1.6285
$pd\sigma$	−2.2675	$pd\pi$	2.4736
$dd\sigma$	−1.5424	$dd\pi$	3.6059
$dd\delta$	−1.7157	$\Delta/3$	0.0195

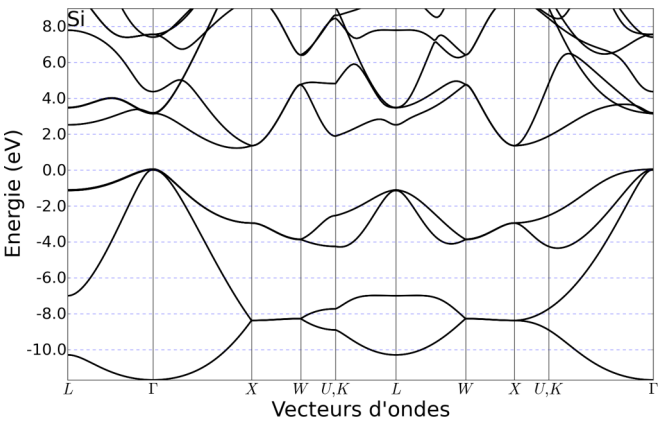

FIG. 9. (Color online) Calculated bulk silicon band structure.

TABLE VI. Some calculated Si band parameters compared with available experimental data.

	TB	Expt.
$\Delta_{\min}$	$0.85\,X$	$0.85\,X$
$E_c(\Delta_{\min})$	1.17 eV	1.17 eV
$m_t(\Delta_{\min})$	0.19	0.19
$m_l(\Delta_{\min})$	0.99	0.98
γ_1	4.5	4.3
γ_2	0.2	0.3
γ_3	1.5	1.6

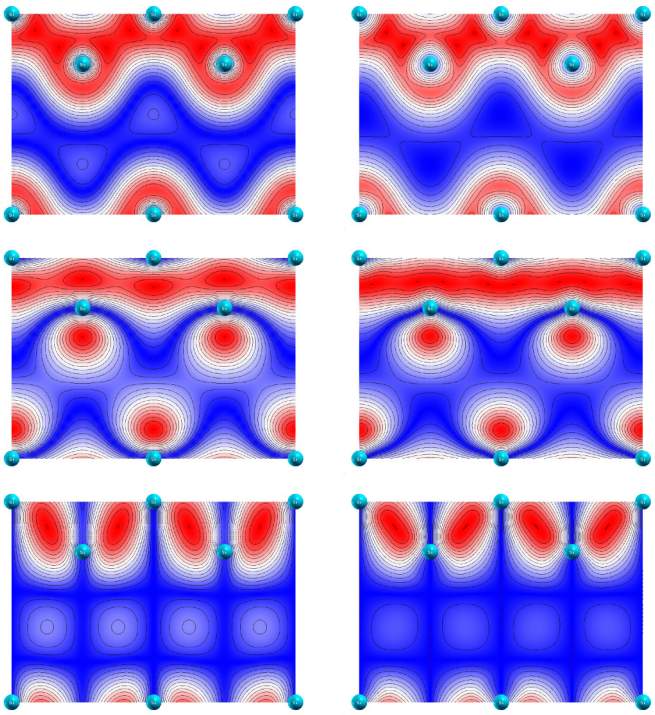

FIG. 10. (Color online) Isodensity contours in the (110) plane for the s_v, z_v, and y'_v conduction Bloch functions at the zone center in bulk Si TB calculation (left) and ABINIT calculations (right).

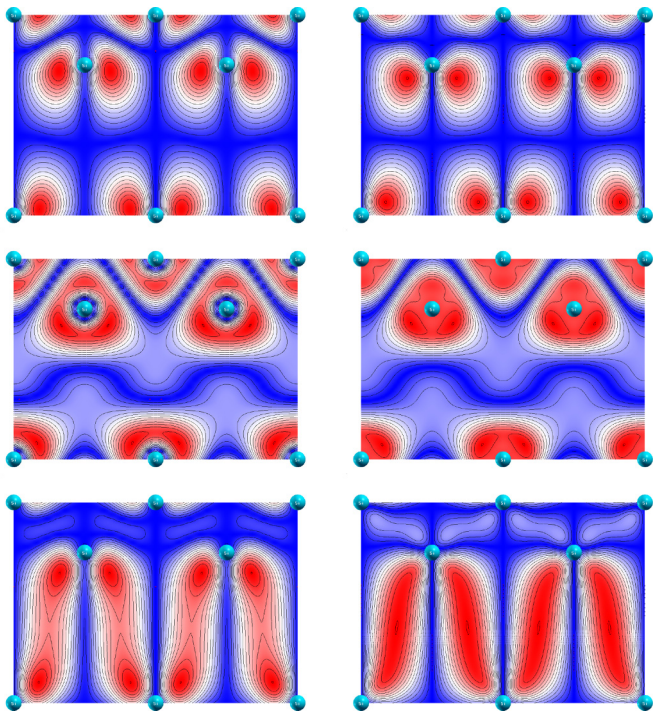

FIG. 11. (Color online) Isodensity contours in the (110) plane for the y'_c, s_c and $d(\Gamma_{12})$ conduction Bloch functions at the zone center in bulk Si TB calculation (left) and ABINIT calculations (right).

vicinity of atomic sites. To clarify this issue, we calculated the electronic wave function using the SIESTA code, which is based on DFT expanded in a set of strictly localized orbitals. In Fig. 8 we show SIESTA results for s_c compared with tight-binding and ABINIT calculations. Tight-binding wave functions compare more favorably with SIESTA results.

APPENDIX B: SILICON PARAMETRIZATION AND SCREENING PARAMETERS

Silicon is technologically the most important semiconductor, and its band structure has been determined experimentally with very high accuracy. Yet, the ability of simple tight-binding models for precise modeling of Si electronic properties has been much discussed [32]. The $spds^*$ model has solved these issues and the parametrization given in Table V provides the band structure shown in Fig. 9, which is in very good agreement with experiment [33] (see Table VI).

In the present context, Si is interesting both for its band structure specificities and because the first quantum numbers n differ from the Ge and GaAs cases. We applied the procedure described in the article to silicon. The final set of screening parameters for Si are $\alpha_s = 1.83$ Å^{-1}, $\alpha_p = 0.97$ Å^{-1}, $\alpha_d = 0.57$ Å^{-1} and $\alpha_{s^*} = 0.72$ Å^{-1}. Figures 10 and 11 show that tight-binding wave functions also compare favorably with ABINIT results in the Si case.

[1] P. W. Anderson, Phys. Rev. **109**, 1492 (1958).
[2] P. W. Anderson, Phys. Rev. **124**, 41 (1961).
[3] J. Hubbard, R. Soc. London Proc. Ser. A **276**, 238 (1963).
[4] P. R. Wallace, Phys. Rev. **71**, 622 (1947).
[5] L. Fu, C. L. Kane, and E. J. Mele, Phys. Rev. Lett. **98**, 106803 (2007).
[6] R. Saito, M. Fujita, G. Dresselhaus, and M. S. Dresselhaus, Phys. Rev. B **46**, 1804 (1992).
[7] J.-C. Slater and G. F. Koster, Phys. Rev. **94**, 1498 (1954).
[8] P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids **44**, 365 (1983).
[9] J. N. Schulman and T. C. McGill, Phys. Rev. Lett. **39**, 1680 (1977).
[10] J.-M. Jancu, R. Scholz, F. Beltram, and F. Bassani, Phys. Rev. B **57**, 6493 (1998).
[11] C. Barreteau, D. Spanjaard, and M. C. Desjonquères, Phys. Rev. B **58**, 9721 (1998).
[12] T. B. Boykin and P. Vogl, Phys. Rev. B **65**, 035202 (2001).
[13] B. A. Foreman, Phys. Rev. B **66**, 165212 (2002).
[14] One should however mention that the question of on-site transition matrix elements, not included in this approach, is still debated within the TB community.

[15] S. Schulz, S. Schumacher, and G. Czycholl, Phys. Rev. B **73**, 245327 (2006).
[16] M. Zieliński, M. Korkusiński, and P. Hawrylak, Phys. Rev. B **81**, 085301 (2010).
[17] J.-C. Slater, Phys. Rev. **36**, 57 (1930).
[18] P.-O. Löwdin, J. Chem. Phys. **18**, 365 (1950).
[19] Note that the orthogonalization procedure induces intra-atomic matrix elements of the momentum operator, compatible with the gauge invariance as discussed by T. Sandu, Phys. Rev. B **72**, 125105 (2005). This, however, does not solve completely the issue of on-site optical matrix elements.
[20] Note that the "Hamiltonian derivation" value of P_{I} also differs from that ($\tilde{P}_{\mathrm{I}} = 0.41$ eV) obtained when fitting a 14-band $\mathbf{k}\cdot\mathbf{p}$ model to tight binding dispersion. $\tilde{P}_{\mathrm{I}}$ corresponds to a renormalized value acting in a restricted basis; see J.-M. Jancu, R. Scholz, E. A. de Andrada e Silva, and G. C. La Rocca, Phys. Rev. B **72**, 193201 (2005).
[21] X. Gonze *et al.*, Comput. Phys. Commun. **180**, 2582 (2009).
[22] X. Gonze *et al.*, Z. Kristallogr. **220**, 552 (2005).
[23] G. L. Bir and G. E. Pikus, *Symmetry and Strain-Induced Effects in Semiconductors* (Wiley, New York, 1975).
[24] S. V. Goupalov and E. L. Ivchenko, Phys. Solid State **43**, 1867 (2001).
[25] J. D. Jackson, *Classical Electrodynamics* (John Wiley, New York, 1975).
[26] R. Nozawa, J. Math. Phys. **7**, 1841 (1966).
[27] R. Resta, Phys. Rev. B **16**, 2717 (1977).
[28] H. Fu, L.-W. Wang, and A. Zunger, Phys. Rev. B **59**, 5568 (1999).
[29] M. Rohlfing and S. G. Louie, Phys. Rev. Lett. **81**, 2312 (1998).
[30] M. Rohlfing and S. G. Louie, Phys. Rev. B **62**, 4927 (2000).
[31] W. Ekardt, K. Losch, and D. Bimberg, Phys. Rev. B **20**, 3303 (1979).
[32] A. S. Martins, T. B. Boykin, G. Klimeck, and B. Koiller, Phys. Rev. B **72**, 193204 (2005).
[33] F. Sacconi, A. Di Carlo, P. Lugli, M. Stadele, and J. Jancu, IEEE Trans. Electron Devices **51**, 741 (2004).