

Extracting E versus \vec{k} effective band structure from supercell calculations on alloys and impuritiesVoicu Popescu^{1,2,*} and Alex Zunger^{3,†}¹*Physics Department and REMRSEC, Colorado School of Mines, Golden, Colorado 80401, USA*²*National Renewable Energy Laboratory, Golden, Colorado 80401, USA*³*University of Colorado, Boulder, Colorado 80302, USA*

(Received 21 October 2011; published 2 February 2012)

The supercell approach to defects and alloys has circumvented the limitations of those methods that insist on using artificially high symmetry, yet this step usually comes at the cost of abandoning the language of E versus \vec{k} band dispersion. Here we describe a computational method that maps the energy eigenvalues obtained from large supercell calculations into an effective band structure (EBS) and recovers an approximate $E(\vec{k})$ for alloys. Making use of supercells allows one to model a random alloy $A_{1-x}B_xC$ by occupying the sites A and B via a coin-toss procedure, affording many different local environments (polymorphic description) to occur. We present the formalism and implementation details of the method and apply it to study the evolution of the impurity band appearing in the dilute GaN:P alloy. We go beyond the perfectly random case, realizing that many alloys may have nonrandom microstructures, and investigate how their formation is reflected in the EBS. It turns out that the EBS is extremely sensitive in determining the critical disorder level for which delocalized states start to appear in the intermediate band. In addition, the EBS allows us to identify the role played by atomic relaxation in the positioning of the impurity levels.

DOI: [10.1103/PhysRevB.85.085201](https://doi.org/10.1103/PhysRevB.85.085201)

PACS number(s): 71.23.An, 71.55.Eq

I. INTRODUCTION

The formation of configurationally disordered alloys $A_{1-x}B_xC$ from ordered constituent solids AC and BC is inevitably associated with loss of long-range order and hence with the automatic loss of the concepts of E versus \vec{k} band structure and its derived quantities such as effective-mass $m^* = \hbar^2[\partial^2 E/\partial k_\alpha \partial k_\beta]^{-1}$, band-velocity $v_{\vec{k}} = \hbar^{-1}\partial E/\partial \vec{k}$, and van Hove singularities. The language used for the description of disordered alloys has thus naturally shifted to wave-vectorless, integrated constructs such as total or local density of states. At the same time, the band structure language of $E(\vec{k})$ continues to be enormously useful to phenomenologically describe trends with alloy composition x from the ordered $x = 0.0$ and $x = 1.0$ constituents to disordered structures for intermediate x . Such descriptions applying heuristically the language of band theory for alloys include the use of effective-mass,^{1,2} van Hove singularities in reflectivity,³ effective band dispersion measured in angular-resolved photoemission,⁴ and magnetotunneling spectroscopy.⁵

We have developed a computational method that maps the energy eigenvalues obtained from large supercell calculations into an effective band structure (EBS).⁶ The supercell may be constructed to represent either a perfectly random $A_{1-x}B_xC$ alloy, or a perfectly ordered ABC compound, or any other intermediate state. The eigensolutions are first obtained by explicitly diagonalizing the supercell single-particle Schrödinger equation and then transforming the results to an EBS. In such a description, each of the C_i atoms located on a lattice site $i = 1, \dots, n_C$ may have a distinctly different local environment (depending on the coordination of the atom C_i by different amounts of A and B atoms) and so the ensuing EBS would correspond to a polymorphic description of the alloy, different from simplified descriptions in alloy theory (e.g., effective medium approximations) where all atoms, at a fixed composition x , are assumed to have the same potential.

In the present work we go beyond the perfectly random alloy, realizing that many alloys have nonrandom microstructures, and ask how is their formation reflected in the effective alloy band structure. Indeed, a supercell can be readily constructed in such a way as to include particular realizations such as total or partial layer ordering, chain formation, and/or clustering. Focusing on the effective $E(\vec{k})$ is different than the more conventional alloy descriptions that rely mostly on density of states. Here we will witness the degree to which the EBS of a (nonrandom) alloy either maintains or loses the sharpness akin to perfect solids or gains or loses it with respect to the completely random system.

To study the manifestation of alloy microstructure on the EBS we have chosen GaN with small amounts of phosphorus alloyed into it. Whereas in the complimentary system GaP:N small amounts of N additions create states that are about $\simeq 300$ meV below the host crystal conduction band minimum,^{7,8} in GaN:P one observes instead deep, localized midgap levels,⁹ developing into an “impurity band.”⁶ Here we follow the dispersion of such an impurity band, as well as the effects of atomic aggregation of phosphorus atoms—random versus clustered—on the EBS.

The impurity band is a generic term denoting a continuous distribution of single-particle energy levels coming together from a collection of impurities with increased concentration.¹⁰ It is, as such, different from the common energy band of an ordered crystal in that the wave vector is no longer a good quantum number.¹¹ Amongst the mechanisms leading to the formation of an impurity band, the Anderson model¹² requires only a certain degree of disorder being present to produce electron localization. An interesting aspect related to the Anderson model is the appearance of a mobility edge upon the formation of the impurity band from a random set of impurities. This is a demarcation energy that separates the localized and delocalized states into different regions of the impurity band. As the impurity atoms become spatially

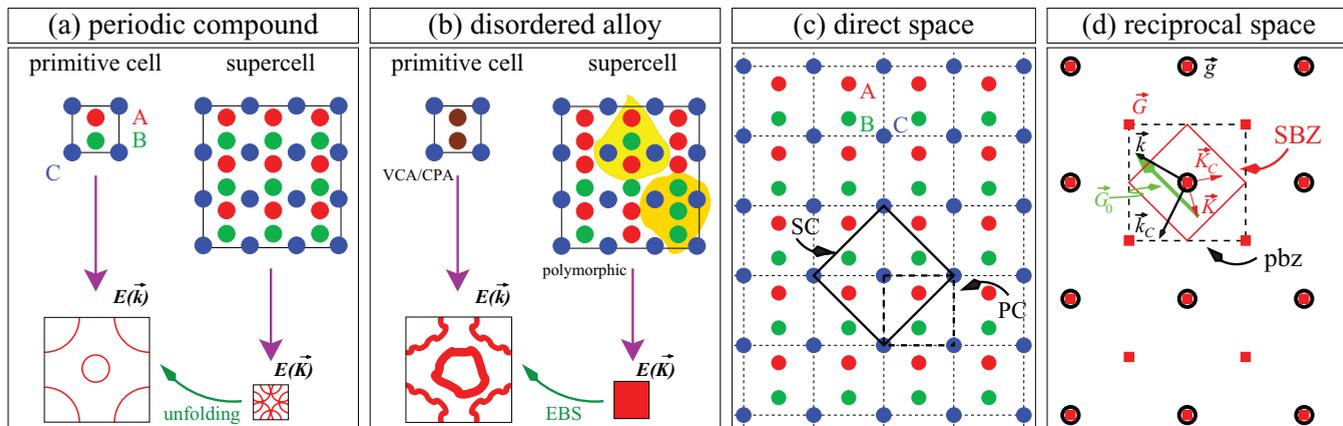


FIG. 1. (Color online) Two-dimensional representation of (a) a simple unfolding of bands for a periodic fictitious system ABC and (b) the analogy with periodic systems when treating an alloy $A_{1-x}B_xC$ by effective medium and polymorphic model approaches. The connection between the two representations of a primitive cell (PC) and a supercell (SC) in (c) direct space and (d) reciprocal space. The PC and SC are related by a transformation similar to Eq. (1). The corresponding primitive (pbz) and supercell (SBZ) Brillouin zones, their associated wave vectors \vec{k} (\vec{K}) and translation vectors \vec{g} (\vec{G}) as well as folding (unfolding) relations are also illustrated. In panel (d), \vec{k}_C and \vec{K}_C are wave vectors equivalent to \vec{k} and \vec{K} by a symmetry operation C of both pbz and SBZ.

ordered and the degree of disorder diminishes, the width of the delocalized impurity states region vanishes. We have found that the impurity band appearing in GaN:P exhibits similar characteristics to the Anderson model: the center of the band is an energy interval corresponding to extended states, whereas below and above it one finds localized levels originating from P–P pairs and P–P–P triplets of various separations.

The impurity band may have important technological applications. For example, if it is located in the gap of an absorber, the impurity band can become the active state in the intermediate band solar cell (IBSC)¹³ as shown recently, for example, in the case of GaAs:N.¹⁴ For an IBSC operation, however, it is of particular importance that the delocalized impurity band states form at the expense of discrete, localized levels that would act as nonradiative recombination centers. We study here, using the EBS technique, precisely this evolution of the impurity band, comparing the perfectly random alloy with systems of increased aggregation of P atoms.

Simple effective medium approaches are not well suited to study the all-important effect of sublattice relaxation, as they enforce an artificially high symmetry. In contrast, our method to determine the alloy EBS takes into account explicitly the local atomic relaxation within the supercell. We investigate the effect of neglecting this relaxation step and find that, in the particular case of GaN:P, it plays the most important role in positioning the impurity band within the GaN band gap. The EBS analysis turns out to be also in this case a useful tool for tracking such manifestations.

The paper is divided into two main parts: we first provide in Sec. II the formalism underlying the EBS determination, outlining the basic principles of band unfolding. We then give in detail all the necessary information on actually constructing an EBS for an alloy using a plane-wave basis for the eigenfunctions in Sec. III. The second part of the paper, Sec. IV, describes the application of the EBS to the study of the appearance and evolution of the impurity band in GaN:P. Our results show that an important aspect of the

method is its ability to derive and describe the impurity band from direct calculations, rather than modeling. As such, the EBS proves to be a useful tool in making the link between extensive supercell calculations and a straightforward, ready interpretation in terms of standard primitive cell band structure.

II. EXTRACTING THE SPECTRAL FUNCTION FROM SUPERCCELL CALCULATIONS

A supercell (SC) is an artificial mathematical construction obtained by stacking a primitive cell (PC) along one or more spatial directions. The PC building block of the SC has to be understood only as lattice (basis vectors and atomic positions) and not as crystal structure (lattice vectors and atomic basis). In this context, the SC has long been used in those electronic structure calculation methods that rely on periodic boundary conditions.^{15–17} Since the concept of the SC in terms of site occupancy is arbitrary, this method affords a flexible description of absence of order.

Although the SC energy bands are directly accessible from calculations, in most cases their interpretation is impeded by their complexity. Restoring an E versus \vec{k} picture within a PC is often more convenient. This restoration is accomplished using band unfolding techniques.^{18–22} The idea is illustrated in Figs. 1(a) and 1(b) for a fictitious two-dimensional (2D) system ABC. When dealing with the ordered, periodic compound ABC [Fig. 1(a)], one can use either a SC or a PC to calculate its electronic structure. The reciprocal space can be described by wave vectors belonging either to the primitive (\vec{k}) or the supercell (\vec{K}) Brillouin zone. In both cases, dispersion relations $E(\vec{k})$ or $E(\vec{K})$ can be obtained directly. An unfolding technique will enable the reconstruction of $E(\vec{k})$ from a directly calculated $E(\vec{K})$. Such a procedure might prove useful in performing band-structure analyses when SCs are used in dealing with impurities or for complex systems with large unit cells.²¹

By analogy with the ordered system, one can also treat the disordered $A_{1-x}B_xC$ alloy [Fig. 1(b)] in two ways. Applying

effective medium theories such as the virtual crystal (VCA)²³ or the single-site coherent potential approximation (CPA)^{24,25} forces a PC periodicity and obtains alloy dispersion relations. In the CPA case, these are given as “broadened bands” in terms of Bloch spectral functions.²⁵ Alternatively, the random system $A_{1-x}B_xC$ can be modeled by using large SCs, in which the atomic sites are occupied by A and B atoms following a coin-toss or other disorder procedures commensurate with the composition x . This leads to the natural occurrence of different local atomic environments around the various atomic sites, as illustrated by the shaded areas in Fig. 1(b), a generic construction termed the *polymorphic model*. The directly calculated SC eigenstates are then projected on a reference Hamiltonian defined over the PC. Following an unfolding step, this will determine an alloy effective band structure (EBS) $E(\vec{k})$. The constructed EBS will resemble a “broadened band complex” only inasmuch as such a picture, of a PC-periodic system, is preserved by the polymorphic model.

The crux of the EBS determination method is represented by the calculation of the spectral weight of a large number²² of SC eigenvalues and the construction of a spectral function. Applications of such techniques made use either of localized^{18,20,21} or plane-wave bases^{19,26} and, in most cases, focused on the spectral decomposition at only a few numbers of (high-symmetry) points in the Brillouin zone. Combining the spectral decomposition with a \vec{k} -unfolding algorithm,²⁰ we have shown⁶ that the EBS of alloys can, in practice, be obtained for any primitive \vec{k} vector.

The procedure of extracting the EBS from SC calculations is versatile and not limited to alloys. While modern computer capabilities allow SC calculations to be performed routinely, access to a PC-related, and thus simpler, $E(\vec{k})$ picture shall provide a useful complementary analysis tool.²¹ For this reason we give in this section the details needed for its implementation, making no specific reference to a particular SC construction and occupation. We focus on the k -folding and \vec{K} -unfolding, showing that, at any choice of a SC/PC combination, all the necessary information is elegantly comprised in the transformation matrix between the two bases.

A. Supercell definition and notations used

One can usually see the SC as a stacking along all or some of the three spatial directions of the PC of a Bravais lattice. The PC lattice vectors \vec{a}_i ($i = 1, 2, 3$) make up the building unit for the SC vectors \vec{A}_i . Here and in the following we denote by small (capital) symbols quantities referring to the PC (SC). In matrix notation, the two sets of basis vectors are related by $\vec{A} = \underline{M} \cdot \vec{a}$, or

$$\begin{pmatrix} \vec{A}_1 \\ \vec{A}_2 \\ \vec{A}_3 \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{pmatrix} \cdot \begin{pmatrix} \vec{a}_1 \\ \vec{a}_2 \\ \vec{a}_3 \end{pmatrix}, \quad m_{ij} \in \mathbb{Z}, \quad (1)$$

where the only condition imposed on the transformation matrix \underline{M} is to be invertible (nonsingular). In the most general case, \underline{M} does not need to be diagonal, that is, the SC and PC unit vectors do not need to be collinear. The elements of \underline{M} are integers ($m_{ij} \in \mathbb{Z}$), which corresponds to the case of a SC commensurate²⁰ to the PC, the only one considered here. An

example is given in Fig. 1(c) for the fictitious 2D system ABC. The volumes of the PC and SC unit cells, v_{PC} and V_{SC} , are related by $V_{SC} = v_{PC} \cdot \det(\underline{M})$.

An obvious consequence of this dual description of the direct (real) space by means of a PC and a SC is the existence of two Brillouin zones in the reciprocal space [Fig. 1(d)]. We distinguish between the primitive Brillouin zone (pbz) and the supercell Brillouin zone (SBZ), with the latter having a smaller volume, $V_{SBZ} = v_{pbz} / \det(\underline{M})$. Following the convention adopted above, we denote by \vec{b}_i (\vec{B}_i) the respective unit cell vectors of the pbz (SBZ), constructed in the usual way²⁷: $\vec{b}_i = (2\pi/v_{PC})(\vec{a}_j \times \vec{a}_k)$, and $\vec{B}_i = (2\pi/V_{SC})(\vec{A}_j \times \vec{A}_k)$. A relation similar to Eq. (1) connects the two reciprocal basis vectors:

$$\begin{pmatrix} \vec{B}_1 \\ \vec{B}_2 \\ \vec{B}_3 \end{pmatrix} = \underline{M}^{-1} \cdot \begin{pmatrix} \vec{b}_1 \\ \vec{b}_2 \\ \vec{b}_3 \end{pmatrix}, \quad (2)$$

emphasizing the requirement of \underline{M} being invertible and showing that $\vec{B}_i \parallel \vec{b}_i$ if and only if \underline{M} is diagonal.

The reciprocal lattice vectors \vec{g}_n (\vec{G}_n) associated with the pbz (SBZ),

$$\vec{g}_n = \sum_i p_i \vec{b}_i, \quad p_i \in \mathbb{Z}, \quad (3a)$$

$$\vec{G}_n = \sum_i q_i \vec{B}_i, \quad q_i \in \mathbb{Z}, \quad (3b)$$

will form two infinite sets $\{\vec{g}_n\}$ and $\{\vec{G}_n\}$ with the obvious property [see also Fig. 1(d)] $\{\vec{g}_n\} \subset \{\vec{G}_n\}$; in other words, any \vec{g} vector is a \vec{G} vector.

B. Folding and unfolding of wave vectors

Folding of states of different wave vectors in supercells depends only on the geometry and symmetry of the SC and its underlying PC, through their corresponding Brillouin zones SBZ and pbz. A wave vector \vec{k} (in pbz) is said to *fold* into a wave vector \vec{K} (in SBZ) [see Fig. 1(d)] if there exists a reciprocal lattice vector \vec{G}_0 such that

$$\vec{K} = \vec{k} - \vec{G}_0. \quad (4)$$

Conversely, a wave vector \vec{K} (of the SBZ) *unfolds* into $\vec{k}_i \in \text{pbz}$ if

$$\vec{k}_i = \vec{K} + \vec{G}_i, \quad i = 1, \dots, N_{\vec{K}}. \quad (5)$$

Despite their apparent equivalence, Eqs. (4) and (5) have been intentionally written down explicitly because they summarize the very principle of folding and unfolding of states. Indeed, the vectors \vec{K} and \vec{G}_0 in Eq. (4) are *unique* for a given \vec{k} , which means that a given wave vector $\vec{k} \in \text{pbz}$ is mapped precisely into a single wave vector $\vec{K} \in \text{SBZ}$ (folding). In contrast, Eq. (5) shows that a given wave vector \vec{K} can be obtained from a number $N_{\vec{K}}$ of different (\vec{k}_i, \vec{G}_i) pairs (unfolding). Not surprisingly,

$$N_{\vec{K}} = \det \underline{M}, \quad (6)$$

which, as we could see, equals v_{pbz} / V_{SBZ} . This is illustrated in Fig. 1(d) for the 2D model system ABC. In this case $N_{\vec{K}} = 2$,

with \vec{K} unfolding into $\vec{k}_1 = \vec{K}$ ($\vec{G}_1 \equiv \vec{0}$) and $\vec{k}_2 = \vec{k}$ ($\vec{G}_2 \equiv \vec{G}_0$). Let us also note here that $N_{\vec{K}}$ refers to the full Brillouin zone, not to its irreducible part; that is, some of the $\vec{k}_i \in \text{pbz}$ wave vectors may be related by symmetry operations of the PC space group.

C. Folding and unfolding of states

Standard electronic structure calculation methods can be applied to a periodic solid using either a PC or a SC representation. Because of periodicity in both PC and SC, \vec{k} (PC) and \vec{K} (SC) are good quantum numbers. By solving the associated Schrödinger equation of the electronic system one can readily obtain both the eigenvectors $|\vec{k}n\rangle$ and $|\vec{K}m\rangle$ (where n and m stand for band indices) and the dispersion relations $E(\vec{k})$ and $E(\vec{K})$, which are well-defined quantities in both representations. The zone folding and unfolding geometric relations lead to the property that any SC eigenvector $|\vec{K}m\rangle$ can be expressed as a linear combination of PC eigenvectors^{19,28} $|\vec{k}_i n\rangle$ ($i = 1, \dots, N_{\vec{K}}$). Commonly termed as *folding* of the bulk pbz into the SBZ, this is formally expressed as

$$|\vec{K}m\rangle = \sum_{i=1}^{N_{\vec{K}}} \sum_n F(\vec{k}_i, n; \vec{K}, m) |\vec{k}_i n\rangle. \quad (7)$$

The purpose of an *unfolding* procedure is to recover, from the SC calculation alone, either (i) the PC eigenvectors $|\vec{k}_i n\rangle$ and their contributions $F(\vec{k}_i, n; \vec{K}, m)$ to the SC eigenstates $|\vec{K}m\rangle$, or (ii) as illustrated in Fig. 1(a), the $E(\vec{k})$ picture from the often complicated $E(\vec{K})$. This last step can be accomplished by projecting $|\vec{K}m\rangle$ on all the PC Bloch states $|\vec{k}_i n\rangle$ of a fixed \vec{k}_i and calculating the spectral weight,^{18,19}

$$P_{\vec{K}m}(\vec{k}_i) = \sum_n |(\vec{K}m|\vec{k}_i n)|^2. \quad (8)$$

This quantity represents the probability of finding a set of PC states $|\vec{k}_i n\rangle$ contributing to the SC state $|\vec{K}m\rangle$, or, equivalently, the amount of Bloch character \vec{k}_i preserved in $|\vec{K}m\rangle$ at the same energy $E_n = E_m$.^{18,19} From $P_{\vec{K}m}(\vec{k}_i)$ one can further derive a spectral function (SF) of continuous variable E :

$$A(\vec{k}_i, E) = \sum_m P_{\vec{K}m}(\vec{k}_i) \delta(E_m - E). \quad (9)$$

In the following section we give examples that illustrate how the SF $A(\vec{k}_i, E)$ is used to unfold a SC-calculated $E(\vec{K})$ into a PC-adapted $E(\vec{k})$.

D. Examples of unfolding

1. The trivial case: Multiple identical primitive cells

The simplest case of unfolding $E(\vec{K})$ into $E(\vec{k})$ is that of a SC obtained by a spatial repetition of identical PCs, without changing the symmetry of the lattice [as in the example of Fig. 1(a)]. By construction, such a SC is not introducing any additional coupling between the PC states $|\vec{k}_i n\rangle$ and thus

$$P_{\vec{K}m}(\vec{k}_i) = g_n(\vec{k}_i) \delta(E_m - E_n), \quad (10)$$

where $g_n(\vec{k}_i)$ is the (bulk) degeneracy of the PC state $|\vec{k}_i n\rangle$ at E_n . The SF will then be a set of δ functions of integer amplitude

$g_n(\vec{k}_i)$ at each eigenvalue $E_n(\vec{k}_i)$ of the PC Hamiltonian, which reconstructs exactly the $E(\vec{k})$ provided by a direct PC calculation [Fig. 1(a)].

2. Superlattices and quantum wells

We recall that, for the purpose of $E(\vec{K})$ unfolding, the PC building block of the SC is understood only as lattice (basis vectors and atomic positions). Indeed, all the information necessary for unfolding is provided by the geometric relation Eq. (1), with all other ingredients deduced from it. For example, one can calculate directly, using a SC construction, the electronic (band) structure $E(\vec{K})$ of a quantum well and/or a superlattice system A_n/B_m , where A and B are zinc-blende III-V materials.^{29,30} The PC of such a system can be chosen either as the fcc or simple cubic (sc) Bravais lattice, and a dispersion relation $E(\vec{k})$ relative to the corresponding Brillouin zone can be derived. Unlike the trivial situation described above and illustrated in Fig. 1(a), the symmetry of the SC is now different from that of the PC. As a result, the bulk states $|\vec{k}_i n\rangle$ may have different representations in the SC and, in addition to folding, they can also couple one to another, differently than in the bulk. Even if materials A and B are ordered compounds, the spectral weights determined from such a calculation, Eq. (8), may no longer be integer values. One can obtain, instead, nonzero $P_{\vec{K}m}(\vec{k}_i)$ for *different* \vec{k}_i wave vectors at the *same* energy E_m . Thus the SF analysis will reveal the identity and the amount of each of the various $|\vec{k}_i n\rangle$ PC eigenstates that couple to form a SC state. As such, the SF decomposition becomes a useful tool that enables a band-structure analysis in the Brillouin zone of the underlying PC. For example, in semiconductor superlattices, one can (i) determine the origin of a band-gap reduction upon ordering as the direct result of bulk states repulsion or (ii) identify a pseudodirect (Γ point) optical transition as an indirect one (X or L point), occurring because of the X and/or L points (pbz) folding into the Γ point (SBZ).³¹

3. Alloy systems

We focus in this paper on applying the SC construction to the electronic structure calculation of substitutionally disordered alloys. In contrast to effective medium theories, by building a large SC, as illustrated in Fig. 1(b), one accommodates a polymorphic description of the system $A_{1-x}B_xC$, allowing different local environments to appear inside the SC. As shown in Fig. 1(b), one can generate a particular random realization of the fictitious alloy that allows different local environments [highlighted areas of Fig. 1(b)] to appear spontaneously around each of the atoms A, B, and C. After solving for the SC Hamiltonian, one obtains the spectral weight given by Eq. (8) and, through Eq. (9), the corresponding SF. In contrast to the SCs of ordered systems, the spectral weights will no longer be δ functions. In particular, the different local environments and the inelastic scattering in the alloy will be reflected in a *finite width* of $A(\vec{k}, E)$ in *both* arguments \vec{k} and E . With the SF determined for a set of \vec{k} vectors and over a wide range of energies, one can obtain the alloy EBS. Its effective construction is described in detail in the next section.

III. CONSTRUCTING THE EFFECTIVE BAND STRUCTURE OF A RANDOM ALLOY

We describe in this section the actual construction of an EBS for an alloy, using the specific example of $\text{In}_{0.1}\text{Ga}_{0.9}\text{N}$, a system that preserves, to a large extent, a recognizable band structure (“a weakly perturbed alloy”).⁶ The necessary steps for an alloy EBS determination—illustrated in Fig. 2—are the following:

(A) Choosing the SC to be used in modeling the alloy system, deciding on a reference PC and a set of wave vectors $\{\vec{k}_i\}$ over which to construct the EBS. As we discuss in more detail below, this set needs to be extended to include also the additional PC wave vectors that are equivalent by symmetry with \vec{k}_i . We denote this extended set by $\{\vec{k}_j\}$;

(B) Decorating the SC, one random realization at a time;

(C) Relaxing the atomic positions so as to minimize the elastic energy;

(D) Calculating the SC eigenvalues and eigenvectors;

(E) Determining the set $\{A(\vec{k}_j, E)\}$ of SFs for all \vec{k}_j vectors of step (A), and calculating an internal average over those \vec{k} vectors that are equivalent by symmetry with \vec{k}_i , which provides a subset of *averaged*, representative SFs $\{\bar{A}(\vec{k}_i, E)\}$; and

(F) Repeating steps (B)–(E) if different random realizations are used. The statistically averaged SFs at each \vec{k}_i are collected into the EBS. This final product (a typical EBS) is shown in Fig. 2(b) and will serve as a general template for the rest of the results presented in this paper.

A. Initial settings: The supercell, the reference primitive cell, and the set of primitive wave vectors

1. Setting up the supercell

When used to model a partially or completely random substitutional alloy $\text{A}_{1-x}\text{B}_x\text{C}$, the SC needs to be constructed in such a way that its symmetry corresponds to the macroscopic (experimental) symmetry of the alloy and, as such, does not introduce artificial symmetry-enforced coupling of states. On the other hand, by virtue of Eq. (3b), the larger the SC, the more different \vec{G}_i 's are obtained and thus the number $N_{\vec{K}}$ of primitive \vec{k}_i vectors for which SFs can be determined from a single, given \vec{K} [see Eq. (5)], depends on the volume of the SC.

However, the actual size of the SC needs to ensure a reasonable balance between accuracy and computational cost. This aspect has been discussed by Zhang and Wang,²² who compared two different approaches: (i) performing a configurational average over many random realizations of a relatively small sized SC (hundred or thousand atoms), and (ii) calculating a single-shot huge SC (up to a quarter million atoms), relying on the ergodic average over many local environments occurring inside the very same SC. The authors found that, while in certain cases the two methods deliver similar results (for example, the band gap), in those properties related to alloy statistics the equivalence is strongly dependent on both the chosen system and its investigated composition. For the alloys presented here, $\text{In}_x\text{Ga}_{1-x}\text{N}$ and the diluted GaN:P, we found no significant qualitative differences in the EBS obtained from thousand- and few hundred-atom SCs. Detailed calculations have shown, for both systems, that the statistical spread in eigenvalues due to composition fluctuations essentially follows an expected $\sqrt{x(1-x)}$ dependence,³² over the whole composition range for $\text{In}_x\text{Ga}_{1-x}\text{N}$ and within $x \leq 0.1$ for $\text{GaN}_{1-x}\text{P}_x$.

2. Choosing the primitive cell

The next step is to choose a PC that will define the pbz, its associated k -space, and will deal as reference for the EBS. Such a choice of the PC unit vectors \vec{a} is not necessarily unique, since the PC building block of the SC has to be understood only as a lattice and not as a crystal structure. It is sufficient that \vec{a} is related to \vec{A} (SC unit vectors) by a relation equivalent to Eq. (1), with a nonsingular transformation matrix \underline{M} . For example, considering the zinc-blende structure as SC, either the fcc or the sc Bravais lattices can be used as PCs. For the purpose of determining an EBS, one further assumes the existence of a set of “virtual crystal” eigenvectors $\{|\vec{k}n\rangle\}$, with $\vec{k} \in \text{pbz}$. As shown below, not only is the explicit calculation of $|\vec{k}n\rangle$ unnecessary, they do not appear at all in the final expression of $P_{\vec{K}_m}(\vec{k}_j)$. The only condition $|\vec{k}n\rangle$ needs to fulfill is that they form a complete, orthonormal set:

$$\langle \vec{k}n | \vec{k}n' \rangle = \delta_{nn'}. \quad (11)$$

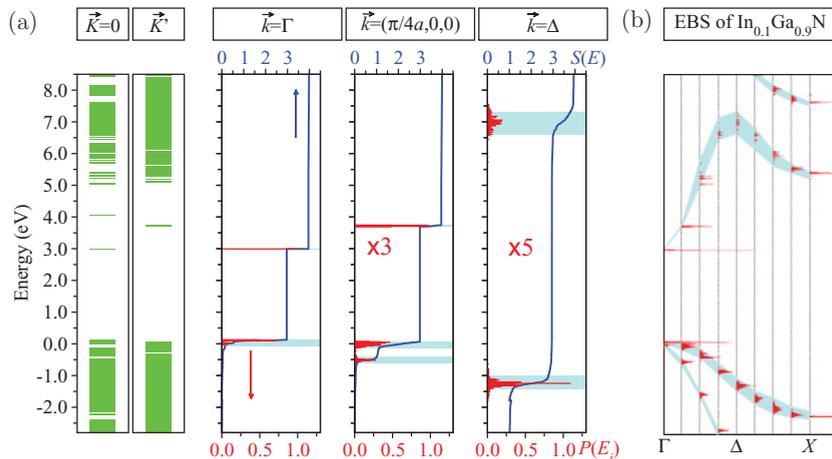


FIG. 2. (Color online) An example, for the $\text{In}_{0.1}\text{Ga}_{0.9}\text{N}$ alloy, of how the effective band structure (EBS) is obtained. (a) Supercell eigenstates $\{E_m(\vec{K}_j), |\vec{K}_j, m\rangle\}$ calculated at various \vec{K}_j are projected and then unfolded on the fcc Brillouin zone vectors $\vec{k}_i = \vec{K}_j + \vec{G}_i$ [Eq. (5)], providing the spectral functions $\bar{A}(\vec{k}_i, E)$ [Eq. (16), red (dark gray)] and the cumulative sums $S_{\vec{k}_i}(E)$ [Eq. (17), blue (black)]; the latter quantity is used to determine the alloy “bands” positions and widths [cyan (light gray) shaded areas]. (b) Results for all \vec{k}_i vectors (thin vertical lines) are put together into an E versus \vec{k} plot, having along the abscissa also the spectral function amplitude.

3. The extended set of primitive and supercell wave vectors

Let us assume that we have decided to construct the alloy EBS for a set $\{\vec{k}_i\}$ of a chosen pbz. We enforce the *macroscopic* symmetry of the alloy by averaging the SFs $A(\vec{k}, E)$ over all symmetry-related \vec{k} vectors. Indeed, if the system were periodic, with C one of its symmetry operations, the PC (SC) wave vectors \vec{k} and $\vec{k}_C = C\vec{k}$ (\vec{K} and $\vec{K}_C = C\vec{k}$) would be equivalent, as shown in Fig. 1(d). For a *random alloy*, however, C is no longer a symmetry operation of the SC, and we therefore need to take into account explicitly all of the corresponding $\vec{K}_C = \vec{k}_C + \vec{G}$ vectors. This procedure increases the size of the chosen set $\{\vec{k}_i\}$, leading to the extended sets $\{\vec{k}_j\}$ and $\{\vec{K}_J\}$ ($J \geq 1$). The latter will provide, for the given SC and PC, all of the primitive vectors \vec{k}_j via Eq. (5).

B. Decorating the supercell

We model a substitutional $A_{1-x}B_xC$ alloy by occupying the unrelaxed lattice sites $\vec{R}_{n\alpha}^0$ with atoms of different species $\alpha = A, B$, and C commensurate with the composition x . This occupation procedure needs to be flexible enough so as to simulate either the completely random or any other intermediate state possessing short- or long-range order.

C. Relaxing the atomic positions

After decorating the SC with a random realization, we allow the atoms to relax about their initial positions $\vec{R}_{n\alpha}^0$ so as to minimize the elastic strain energy in the SC. During this process, the bonds A–C, A–B, and B–C of an $A_{1-x}B_xC$ alloy will tend towards their “natural” (equilibrium) values in the corresponding binaries, leading to local relaxation. We calculate the elastic energy by means of a valence force field (VFF) functional^{33,34} in its generalized form.³⁵ Its expression is given as a sum of three terms, representing the bond stretching, bond bending, and bond-length/bond-angle interactions. Each of these terms is described by VFF parameters which are related to the elastic coefficients of the corresponding bulk materials.³⁵ The importance of this relaxation step for the impurity band appearance and positioning is analyzed in detail in Sec. IV D.

D. Calculating the SC eigenvalues and eigenvectors

At a given random realization, we obtain the SC eigenvalues and eigenvectors by solving the single-particle equation

$$\left[-\frac{\beta}{2} \nabla^2 + \sum_{n,\alpha} \hat{v}_\alpha(\vec{r} - \vec{R}_{n\alpha}, \underline{\varepsilon}_n) + \hat{V}_{\alpha\text{NL}} \right] |\vec{K}m\rangle = E_m |\vec{K}m\rangle, \quad (12)$$

where $v_\alpha(\vec{r} - \vec{R}_{n\alpha}, \underline{\varepsilon}_n)$ is a screened atomic empirical pseudopotential depending on the identity α of the atom and the local strain tensor $\underline{\varepsilon}$ at its relaxed position $\vec{R}_{n\alpha}$:

$$v_\alpha(\vec{r}, \underline{\varepsilon}) = v_\alpha(\vec{r}, 0)[1 + \gamma_\alpha \text{Tr}(\underline{\varepsilon})], \quad (13)$$

with γ_α a fitting parameter introducing a further dependence on the identity of the neighbors.³⁶ The other terms entering Eq. (12) are the nonlocal spin-orbit coupling potential $\hat{V}_{\alpha\text{NL}}$ and a scaling factor β for the kinetic energy.³⁶

Using the pseudopotentials of each atom α and the relaxed positions $\vec{R}_{n\alpha}$, we solve the single-particle equation (12) by making a plane-wave ansatz³⁷ for the eigenvector $|\vec{K}m\rangle$,

$$|\vec{K}m\rangle = \left[\sum_{\vec{G}} C_{\vec{K}m}(\vec{G}) e^{i\vec{G}\vec{r}} \right] e^{i\vec{K}\vec{r}}, \quad \vec{K} \in \text{SBZ}, \quad (14)$$

where \vec{G} are reciprocal lattice vectors in units of the SBZ, as given by Eq. (3b). The numerical determination of $|\vec{K}m\rangle$ and E_m is accomplished by diagonalizing the Hamiltonian using the folded spectrum method.³⁸

We show in Appendix that, adopting a plane-wave expansion for the SC eigenvectors, the spectral weight $P_{\vec{K}m}(\vec{k}_j)$ of Eq. (8) is given by

$$\begin{aligned} P_{\vec{K}m}(\vec{k}_j) &= \sum_{\vec{g}} |C_{\vec{K}m}(\vec{g} + \vec{k}_j - \vec{K})|^2, \\ &= \sum_{\vec{g}} |C_{\vec{K}m}(\vec{g} + \vec{G}_j)|^2, \end{aligned} \quad (15)$$

where we took into account the unfolding relation, Eq. (5). As anticipated, $P_{\vec{K}m}(\vec{k}_j)$ does not require the explicit knowledge and calculation of the projecting Bloch functions $|\vec{k}n\rangle$. We recall here that $\{\vec{g}_j\} \subset \{\vec{G}_j\}$. Thus all $C_{\vec{K}m}(\vec{g} + \vec{G}_j)$ coefficients in Eq. (15) are well-defined quantities. In addition, as any \vec{g} and \vec{G} is obtained according to one of the Eqs. (3a) or (3b), this shows how $P_{\vec{K}m}(\vec{k}_j)$ explicitly depends on the choice of the reference primitive Brillouin zone (pbz) and its relation to the SBZ.

E. Determining the spectral functions

In order to obtain the SFs, Eq. (12) is solved at each previously determined \vec{K}_J (Sec. III A 3) for thousands of SC eigenstates $\{E_m(\vec{K}_J), |\vec{K}_J m\rangle\}$. We then derive the SFs $A(\vec{k}_j, E)$ using Eqs. (9) and (15). In line with the macroscopic symmetry enforcement procedure, the effective SF $\bar{A}(\vec{k}_i, E)$ at a representative point \vec{k}_i is obtained as the average of SFs of all n_C wave vectors \vec{k}_C belonging to the same symmetry class $C(\vec{k}_i)$ as \vec{k}_i :

$$\bar{A}(\vec{k}_i, E) = \frac{1}{n_C} \sum_{\vec{k}_C \in C(\vec{k}_i)} A(\vec{k}_C, E). \quad (16)$$

F. Constructing the EBS

If more than one random realization is used in calculating the EBS, the symmetry-averaged SFs $\bar{A}(\vec{k}_i, E)$ are collected to give the final symmetry and statistically averaged SF at each \vec{k}_i . Such SFs for alloys are thus broadened and structured, exhibiting a finite, \vec{k} -dependent bandwidth, a direct consequence of the polymorphic nature of the adopted description.

The SC eigenvalues calculation and the SF generation are illustrated for the $\text{In}_{0.1}\text{Ga}_{0.9}\text{N}$ alloy in Fig. 2(a), where the SC eigenvalues have been obtained for two different SC wave vectors and the SFs $\bar{A}(\vec{k}_i, E)$ at selected primitive \vec{k}_i vectors are shown in red (dark gray) along the abscissas. A useful quantity

in the EBS construction and discussion is the cumulative sum,²⁰

$$S_{\vec{k}_i}(\mathcal{E}_n) = \int^{\mathcal{E}_n} \bar{A}(\vec{k}_i, E) dE, \quad (17)$$

depicted in the same panels as $\bar{A}(\vec{k}_i, E)$ with blue (black) lines. This cumulative sum is characterized by steps of value $g(\vec{k}_i)$ whenever an ‘‘alloy band’’ of degeneracy $g(\vec{k}_i)$ is crossed,²⁰ and thus allows one to estimate the alloy bands positions and widths. For the present applications, a \vec{k}_i -dependent bandwidth [the cyan (light-gray) shaded areas in Fig. 2(a)] is taken as the energy range in which $\bar{A}(\vec{k}_i, E) \geq 10^{-3}$ around a band center. The final step in obtaining the EBS is collecting all of the SFs and the bandwidths in single E versus \vec{k} plots, as those shown in Fig. 2(b), where the individual \vec{k}_i vectors are designated by vertical thin lines to which the corresponding SFs $\bar{A}(\vec{k}_i, E)$ are aligned. Thus the abscissa of such a plot simultaneously shows two variables, the wave vector \vec{k} and the spectral function amplitude at each \vec{k}_i .

The EBS of $\text{In}_{0.1}\text{Ga}_{0.9}\text{N}$ shown in Fig. 2(b) is a typical example of a weakly perturbed alloy, exhibiting rather well-defined, albeit broadened, alloy bands. This broadening is different for various bands and \vec{k} vectors: (i) Sharp, bandlike peaks in the SFs are obtained only at the center and edges of the fcc Brillouin zone. (ii) For intermediate \vec{k}_i vectors, the Bloch character of the alloy bands is strongly diminished, especially for the first conduction band (CB1), which reaches a broadening of $\simeq 1$ eV. This is mainly due to the convexity of CB1 in both parent binaries.⁶ Indeed, for \vec{k} inside the pbz, there are more eigenstates of different bulk wave vectors, with closer eigenvalues, that are folding in the same small energy range to form the SC (alloy) eigenstates. (iii) In contrast to CB1, the valence bands and the second conduction band (CB2), have narrower \vec{k} -dependent bandwidths. Note that the calculations for (In,Ga)N were done without spin-orbit coupling, and, as such, the heavy hole and light hole bands are not resolved by the EBS, leading to relatively broader widths of the VB. In turn, the results we obtained and present in the following for the strongly perturbed system GaN:P did include the spin-orbit coupling.

IV. THE EMERGENCE OF AN IMPURITY BAND IN DILUTED GaN:P AND ITS EVOLUTION WITH CLUSTERING AND RELAXATION

We are presenting results of an EBS determination method that have been obtained for the highly mismatched zinc-blende $\text{GaN}_{1-x}\text{P}_x$ alloy^{8,39} in the dilute limit, $x \leq 0.05$. This system represents a prototype for a strongly perturbed alloy as it is characterized by a large natural valence band offset [$\Gamma_{15v}(\text{GaP})$ 1.71 eV above $\Gamma_{15v}(\text{GaN})$] and lattice misfit (21%).⁴⁰ Considering the perfectly random case, previous investigations^{6,9} have shown that the presence of P atoms leads to the occurrence of deep, localized midgap levels (t_2 -like) that develop into an impurity band (IB).⁶ We emphasize here that the term ‘‘completely’’ (or ‘‘perfectly’’) random refers only to the occupation of each of the cation/anion sublattices according to random statistics (which has finite probability for clusters). In our model the alloys still preserve a short-range

order, in that the anions (cations) do not occupy the cation (anion) sublattice. We apply the EBS method to study: (i) the changes in the dispersion of the IB under the effect of atomic aggregation of P, going beyond the completely random case; and (ii) how the EBS is reflecting the neglecting of atomic relaxation caused by the huge lattice mismatch between GaN and GaP, a step otherwise always accounted for in our calculations.

A. Calculation of alloy eigenstates

For all our calculations we use a 512-atom cubic SC, the same size used in the perfectly random case.⁶ We construct twelve random realizations at each composition investigated, performing the statistical average of the determined SFs. As a reference PC we use the fcc Bravais lattice with the lattice constant given by the Vegard law. Detailed comparison of the averaged EBS with one obtained from a 4096-atom SC showed no significant quantitative and qualitative differences. In solving the SC Hamiltonian, we use the same unstrained empirical pseudopotentials corresponding to GaN and GaP— $v_\alpha(\vec{r}, 0)$ in Eq. (13)—as used by Mattila *et al.*^{9,41} including spin-orbit coupling.

B. Impurity states in strongly perturbed alloys

While the appearance of an IB in GaN:P has been demonstrated previously for the perfectly random case,⁶ we shall contrast here the EBS of the completely random systems with some in which partial or total clustering of the substitutional impurity is imposed. Indeed, the single impurity levels will merge together to form the IB upon increasing concentration,^{10,11} but a critical degree of impurity disorder is required for delocalized states to appear in the center of the IB. In the perfectly random GaN:P system we have found that the IB is characterized by a central area of high density of *delocalized* states and a series of isolated peaks, corresponding to localized states originating from P–P pairs and P–P–P triplets, above and below it.⁶ While in a disordered alloy such distributions appear randomly with a given probability, one can imagine, for example, the formation of clumps of P atoms that will manifest as tiny quantum dots, giving rise to discrete, strongly localized levels, removing the continuous dispersion of energy levels specific to an IB. Between the limit of a total clustering and that of complete randomness, characterized by an Anderson-like IB, a transition is expected to occur. We can follow this transition by means of the alloy EBS, calculated by modeling different degrees of phosphorus aggregation.

Figure 3 gives results of our calculations for $\text{GaN}_{0.98}\text{P}_{0.02}$, and Fig. 4 for $\text{GaN}_{0.95}\text{P}_{0.05}$. The leftmost panel of these figures shows the EBS of the completely random system, where the energy zero is the GaN valence band maximum (VBM). The formation of an IB with increasing P composition can be readily recognized in these figures.

It also becomes apparent that, in the perfectly random case, the IB exhibits the main characteristics of an Anderson-like impurity band: (i) the maximum in each $\bar{A}(\vec{k}, E)$ appears at the energy center of the IB, with its center of gravity being pinned in energy and varying very little with P composition; (ii) above and below the center of the IB one observes discrete levels that

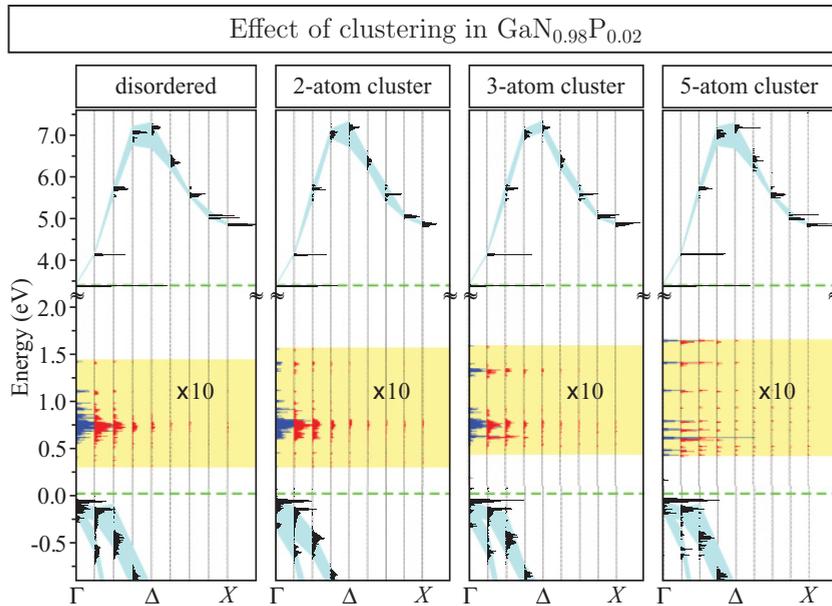


FIG. 3. (Color online) Effect of clustering on EBS for $\text{GaN}_{0.98}\text{P}_{0.02}$ derived from a 512-atom SC. This contains, at the P composition of 0.02, five P atoms in the SC. Leftmost panel: The perfectly random alloy shows the Anderson-like impurity band. Rightmost: All of the five P atoms are grouped in a cluster, leading to discrete levels. In between these extrema we have systems consisting of an n_c -atom P cluster kept fixed while the rest of the P atoms, $5 - n_c$, were placed randomly in the SC (taking 12 different realizations). We note the formation of the impurity band with the decreasing size, from right to left, of the P cluster. The dashed horizontal lines mark the position of the GaN VBM and CBM.

have been identified as stemming from various P-P pairs;⁶ (iii) $\bar{A}(\Gamma, E)$, the SF at the pbz center, is the dominant component, followed by a fast decay for the wave vectors near the zone edge. The increase of P content x_P leads to a rapid widening of the IB, which extends over $\simeq 1$ eV already at $x_P = 0.02$. Simultaneously, this increase also leads to a disintegration of the valence band (VB). Indeed, even at the low composition $x = 0.02$, the VB exhibits widespread \bar{k} -resolved SFs, with a broadening of the EBS in both \bar{k} and E . This effect is much less pronounced in the conduction band (CB), which preserves a weakly perturbed alloy character even at higher P composition.

C. Appearance of discrete levels for clusters

We model the aggregation of phosphorus atoms by constructing spherical clusters of different sizes inside the SC. For the 512-atom cubic SC, there is a number $n_{\max} = [x \times 256]$ of anion positions that P can occupy at a P content x in

$\text{GaN}_{1-x}\text{P}_x$. This number also represents the maximum size of a P cluster we can build in the SC ($[\dots]$ denoting the integer part). We define a cluster of size $n_c \leq n_{\max}$ as the smallest volume, in the center of the SC, occupied by the n_c P atoms distributed exclusively on anion sites. The $n_{\max} - n_c$ atoms left after setting up the cluster are distributed randomly over the rest of the SC; this way, the system we model can be seen as a disordered Ga(N,P) alloy of nominal random composition $(n_{\max} - n_c)/256$ with an n_c -sized P cluster in it. Also, for these P-aggregated systems a number of 12 different realizations were calculated and averaged (except, of course, for $n_c = n_{\max}$) to give the alloy EBS. In contrast to the perfectly random situation, however, the various realizations have the cluster region identical.

In the rightmost panel of Figs. 3 and 4 is shown the EBS for the systems with a cluster of maximum size $n_c = n_{\max}$, for $\text{GaN}_{0.98}\text{P}_{0.02}$ ($n_{\max} = 5$) and $\text{GaN}_{0.95}\text{P}_{0.05}$ ($n_{\max} = 13$), respectively, at the corresponding composition. Although we

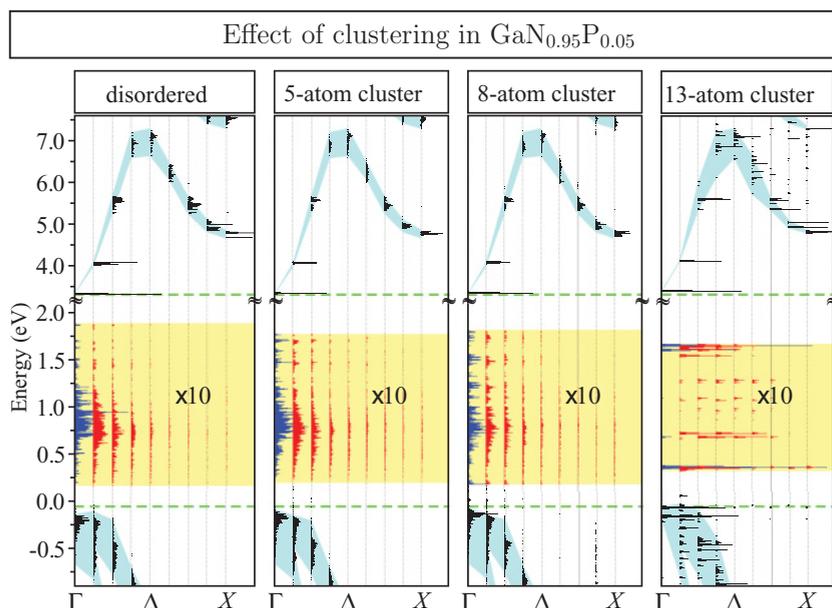


FIG. 4. (Color online) Same as Fig. 3 but for $\text{GaN}_{0.95}\text{P}_{0.05}$. The SC contains at the P composition of 0.05 13 P atoms. Leftmost panel: The perfectly random alloy shows the Anderson-like impurity band. Rightmost: All of the 13 P atoms are grouped in a cluster, leading to discrete levels.

have performed calculations for all possible n_c values between the two limits—total randomness ($n_c = 1$) and complete clustering ($n_c = n_{\max}$)—we show only two illustrative cases, in which the P clusters are of intermediate size $1 < n_c \leq n_{\max}$. Here we should note that, when modeling a disordered $A_{1-x}B_x$ system within a SC approach the composition x takes on only discrete values, rather than being a continuous variable.

The general trend to be observed when moving from perfect disorder ($n_c = 1$) to complete clustering ($n_c = n_{\max}$) is the fragmentation of the IB from a relatively compact band, with a high DOS area in its center, to a collection of discrete levels, akin those of a tiny quantum dot. The VB and CB, on the other hand, show nearly no modification upon clustering, neither in structure, nor in broadening.

The two compositions $x = 0.02$ and $x = 0.05$ exhibit a quite different behavior. In the low composition range ($x = 0.02$) the isolated high-energy peaks (at $\simeq 1.2$ – 1.4 eV) in the EBS originate from localized P-dimer and P-trimer states. Upon clustering, $n_c = 2$ and $n_c = 3$, a satellite feature develops in this high-energy region, separated from the rest of the IB. The IB survivor remains relatively well defined, centered around $E = 0.7$ eV. Details of the EBS, showing only $\bar{A}(\Gamma, E)$ (the dominant component at the pbz center), are plotted in Fig. 5. In this figure the evolution from a rather compact IB (disordered system) to the collection of single localized levels (5-atom cluster) is even easier to follow. The width of the delocalized region of the IB (as long as present, for $n_c \leq 3$) does not change significantly from about 0.45 eV, its value for the completely random case.

For the higher composition, $x = 0.05$, even a high degree of clustering is not sufficient to destroy the delocalized region of the IB, as can be seen for the system with $n_c = 8$ shown in Fig. 4. This is not surprising, since the nominal random P composition outside the cluster region is slightly above 0.02, which would correspond to the previous (low-composition) totally random case. In fact, based on a detailed analysis for all the $1 \leq n_c \leq n_{\max}$ possible cluster sizes (not shown), we could determine that values above 1% for the nominal host randomness are sufficient to ensure the appearance of delocalized states in the IB.

D. Effect of atomic displacements on the EBS

In order to investigate the effect of internal relaxation on the EBS, we do an additional eigenvalues/SF calculation immediately after setting up the SC, *without* relaxing the atomic positions, that is, skipping step (C) in the EBS construction (Sec. III). This set of results will be denoted as “unrelaxed” in the following and will be compared with the EBS obtained following the standard procedure, including relaxation (denoted “relaxed”). We discuss in the following the relaxation effect for completely random GaN:P comparing a single realization (with and without relaxation), having found similar changes for all configurations considered.

It was recognized quite early that, in the case of an isoelectronic impurity, the deep levels occurring in the gap of the host are caused by an attractive potential induced by the local relaxation around the impurity.⁷ Using a similar polymorphic model, Kent and Zunger⁸ have shown that this is indeed the case for the nitrogen impurity level in GaP. This

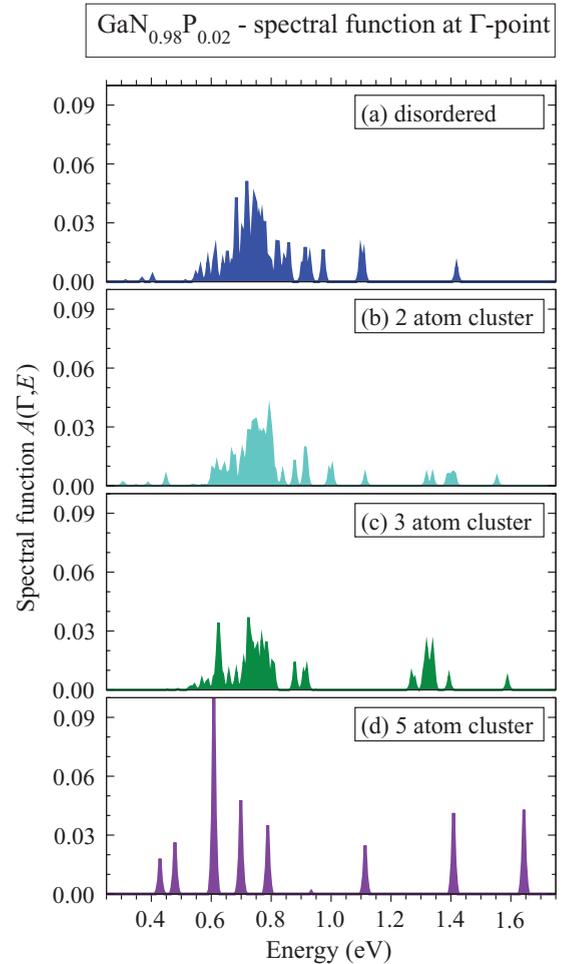


FIG. 5. (Color online) Detailed representation of the spectral function $\bar{A}(\Gamma, E)$ at the Γ point in the energy range of the impurity band for the systems of Fig. 3.

level, appearing about 300 meV below the GaP CBM when proper relaxation is taken into account, is resonant with the host CB as soon as relaxation is ignored. The relaxation around the impurity is the main cause for the P impurity levels appearing deep in the gap of the GaN host.⁶ This is demonstrated in Fig. 6 for various P compositions x , by comparing side by side the EBS of single realizations before and after the internal relaxation.

We observe that, regardless of x , the impurity level (band, for higher x) completely disappears from the gap in the unrelaxed systems. In addition, the GaP hole states are pushed very far below the GaN VBM and thus the valence alloy bands of the unrelaxed GaN:P are narrower than in the relaxed case. The CB, however, suffers a large broadening when relaxation is ignored. This broadening is particularly strong in the vicinity of the X point, where the lowest GaP electron states are pushed down and, as a result of folding, mix with the GaN CB states.

These modifications in the relative positioning of the various states are easy to understand considering the two effects occurring as a result of P substituting for N. First, there is a difference in the electronegativity of the two anions, which is described by the atomic pseudopotentials. Second, the Ga–P bonds formed around the P impurity are subjected to a huge

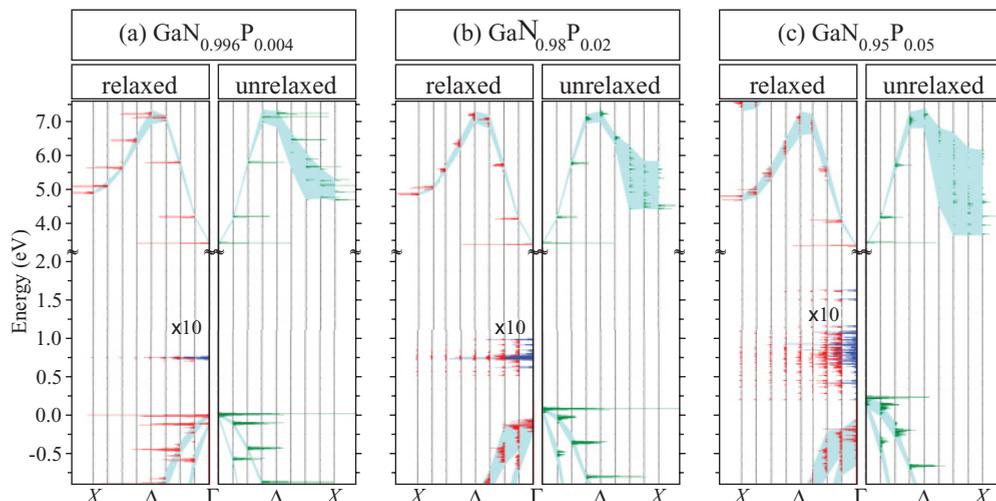


FIG. 6. (Color online) Effect of VFF relaxation on the EBS of GaN:P for different P compositions ($x = 0.004, 0.02, 0.05$) (from left to right) and a fixed random realization. When the relaxation is neglected (the panels labeled “unrelaxed”) the impurity band disappears from the gap.

compression when accommodating the much smaller (21%) GaN lattice constant. The model calculations remove the inherent relaxation step that would naturally occur. The two-anion electronegativity difference alone, although significantly changing the EBS of the GaN:P as compared to GaN, is not sufficiently strong to position the P levels deep into the gap. Indeed, taking the GaN VBM (Γ_{15v}) as reference energy, the natural band edges are, respectively, $\Gamma_{1c}(\text{GaN}) = 3.2$ eV, $\Gamma_{15v}(\text{GaP}) = 1.7$ eV, and $X_{1c}(\text{GaP}) = 4.1$ eV. With the atomic displacements properly taken into account, both Ga–N and Ga–P bonds will tend towards their “natural” lengths and, as a consequence, the corresponding levels will not change too much, resulting in a VB and a CB of the GaN:P alloy being almost exclusively formed by the nearly unstrained GaN bands. If relaxation is ignored, however, Ga–N bonds are practically the same (and so are the respective band edges), but the Ga–P suffer a large contraction. It is this pressure that lowers dramatically the GaP valence and conduction band states, bringing them in resonance with the GaN bands with which they couple, forming wide alloy bands. The former impurity band (originating from the GaP Γ_{15v} level) is now completely buried into the alloy VB. Such a dramatic effect is, of course, an artefact of any calculation which ignores relaxation in highly mismatched systems.

V. CONCLUSIONS

We have presented to a great level of detail a method that maps the energy eigenvalues obtained from large supercell calculations into an effective band structure (EBS) and is able to recover an approximate $E(\vec{k})$ for alloys. The method has been applied to study the evolution of the impurity band appearing in the dilute GaN:P alloy, going beyond the perfectly random case by allowing phosphorus aggregation to occur. We found that the EBS is sensitive enough to determine the critical disorder level for which delocalized states start to appear in

the intermediate band and established that this happens, for GaN:P in the zinc-blende structure, at around 1% phosphorus composition. We have also investigated the role played by the local atomic relaxation in the positioning of the impurity levels and found that the omission of this step leads to the disappearance of these levels from the band gap of the GaN:P alloy. We suggest that the EBS has significant merits to become a link between modern, expensive supercell calculations and a rather simple interpretation of their results in terms of an easily recognizable band structure.

ACKNOWLEDGMENTS

This work was supported by the US Department of Energy, Office of Science, Basic Energy Sciences, under Contract No. DE-AC36-08GO28308 with the National Renewable Energy Laboratory, Golden, Colorado. V.P. also acknowledges the administrative support of REMRSEC at the Colorado School of Mines, Golden, Colorado.

APPENDIX: DERIVATION OF THE SPECTRAL WEIGHT IN A PLANE-WAVE BASIS

This Appendix provides a plane-wave expression for the spectral weight [Eq. (8)],

$$\begin{aligned} P_{\vec{K}m}(\vec{k}_0) &= \sum_n |\langle \vec{K}m | \vec{k}_0 n \rangle|^2 \\ &= \sum_n \langle \vec{K}m | \vec{k}_0 n \rangle \langle \vec{k}_0 n | \vec{K}m \rangle, \end{aligned} \quad (\text{A1})$$

quantifying the amount of a *fixed* Bloch character \vec{k}_0 preserved in the SC eigenvector $|\vec{K}m\rangle$, given, in the present case, by a plane-wave expansion [Eq. (14)]. We follow very general guidelines that should help in establishing analogous expressions for the spectral weight, regardless of the basis functions used to represent the SC eigenvectors.

We assume the PC eigenvectors $|\vec{k}_0 n\rangle$ to be of an analogous form

$$|\vec{k}_0 n\rangle = u_{\vec{k}_0 n}(\vec{r}) e^{i\vec{k}_0 \vec{r}} = \left[\sum_{\vec{g}} B_{\vec{k}_0 n}(\vec{g}) e^{i\vec{g} \vec{r}} \right] e^{i\vec{k}_0 \vec{r}} \quad \vec{k} \in \text{pbz}, \quad (\text{A2})$$

and satisfying the orthogonality condition [Eq. (11)]. We note that the distinction between the two mappings, SBZ and pbz, of the reciprocal space is directly reflected in the two different summations over \vec{G} and \vec{g} in Eqs. (14) and (A2). From Eq. (11), using the ansatz (A2) one obtains the following:

$$\begin{aligned} \delta_{nn'} &= \langle \vec{k}_0 n | \vec{k}_0 n' \rangle \\ &= \sum_{\vec{g}, \vec{g}'} B_{\vec{k}_0 n}^*(\vec{g}) B_{\vec{k}_0 n'}(\vec{g}') \int d^3 r e^{i(\vec{g}' - \vec{g}) \vec{r}} \\ &= \sum_{\vec{g}} B_{\vec{k}_0 n}^*(\vec{g}) B_{\vec{k}_0 n'}(\vec{g}), \end{aligned}$$

or

$$\sum_{\vec{g}} B_{\vec{k}_0 n}^*(\vec{g}) B_{\vec{k}_0 n'}(\vec{g}) = \delta_{nn'}. \quad (\text{A3})$$

One can look upon the expansion coefficients $B_{\vec{k}_0 n}(\vec{g})$ as elements of a matrix $\underline{\tilde{B}}$, with $\tilde{B}_{n\vec{g}} = B_{\vec{k}_0 n}(\vec{g})$. In addition, completeness of the basis set $\{|\vec{k}_0, n\rangle\}$ implies that the number of PC (bulk) states n and that of the plane-waves $e^{i\vec{g} \vec{r}}$ is equal, that is, $\dim\{n\} = \dim\{\vec{g}\}$. Thus the matrix $\underline{\tilde{B}}$ is quadratic and Eq. (A3) can be put into matrix form:

$$\underline{\tilde{B}}^\dagger \cdot \underline{\tilde{B}} = \underline{\mathbb{1}},$$

where $\underline{\mathbb{1}}$ is the unit matrix. Obviously,

$$\underline{\tilde{B}} \cdot \underline{\tilde{B}}^\dagger = \underline{\mathbb{1}}$$

also holds, which is equivalent to

$$\sum_n B_{\vec{k}_0 n}^*(\vec{g}) B_{\vec{k}_0 n}(\vec{g}') = \delta_{\vec{g}, \vec{g}'}. \quad (\text{A4})$$

Recalling that any \vec{g} vector is simultaneously a \vec{G} vector ($\{\vec{g}_i\} \subset \{\vec{G}_i\}$) one can show that any given \vec{G}_0 can be expressed as $\vec{G}_0 = \vec{g}_0 + \vec{k}$ [see also Fig. 1(b)]. Applying this mapping, the SC eigenvector $|\vec{K} m\rangle$ [Eq. (14)] may be written as

$$\begin{aligned} |\vec{K} m\rangle &= \left[\sum_{\vec{G}} C_{\vec{K} m}(\vec{G}) e^{i\vec{G} \vec{r}} \right] e^{i\vec{K} \vec{r}} \\ &= \left[\sum_{\vec{g}} \sum_{\vec{k}} C_{\vec{K} m}(\vec{g} + \vec{k}) e^{i(\vec{g} + \vec{k}) \vec{r}} \right] e^{i\vec{K} \vec{r}} \\ &= \left[\sum_{\vec{k}} e^{i\vec{k} \vec{r}} \sum_{\vec{g}} C_{\vec{K} m}(\vec{g} + \vec{k}) e^{i\vec{g} \vec{r}} \right] e^{i\vec{K} \vec{r}}. \quad (\text{A5}) \end{aligned}$$

We use Eqs. (A2) and (A5) to get

$$\begin{aligned} \langle \vec{k}_0 n | \vec{K} m \rangle &= \sum_{\vec{k}} \sum_{\vec{g}, \vec{g}'} B_{\vec{k}_0 n}^*(\vec{g}) C_{\vec{K} m}(\vec{g}' + \vec{k}) \\ &\quad \times \int d^3 r e^{i(\vec{g}' - \vec{g}) \vec{r}} e^{i(\vec{K} - \vec{k}_0 + \vec{k}) \vec{r}}. \end{aligned}$$

Because \vec{K} , \vec{k}_0 , and \vec{k} are all inside the pbz, the integral of the last equation satisfies

$$\int d^3 r e^{i(\vec{g}' - \vec{g}) \vec{r}} e^{i(\vec{K} - \vec{k}_0 + \vec{k}) \vec{r}} = \delta_{\vec{g}, \vec{g}'} \delta_{\vec{k}, \vec{k}_0 - \vec{K}},$$

and thus

$$\langle \vec{k}_0 n | \vec{K} m \rangle = \sum_{\vec{g}} B_{\vec{k}_0 n}^*(\vec{g}) C_{\vec{K} m}(\vec{g} + \vec{k}_0 - \vec{K}). \quad (\text{A6})$$

Analogously,

$$\langle \vec{K} m | \vec{k}_0 n \rangle = \sum_{\vec{g}} B_{\vec{k}_0 n}(\vec{g}) C_{\vec{K} m}^*(\vec{g} + \vec{k}_0 - \vec{K}). \quad (\text{A7})$$

Inserting the last two equations in Eq. (A1) and making use of Eq. (A4), one arrives at

$$\begin{aligned} P_{\vec{K} m}(\vec{k}_0) &= \sum_{\vec{g}, \vec{g}'} C_{\vec{K} m}(\vec{g} + \vec{k}_0 - \vec{K}) C_{\vec{K} m}^*(\vec{g}' + \vec{k}_0 - \vec{K}) \delta_{\vec{g}, \vec{g}'} \\ &= \sum_{\vec{g}} |C_{\vec{K} m}(\vec{g} + \vec{k}_0 - \vec{K})|^2, \quad (\text{A8}) \end{aligned}$$

which is exactly the sought result [Eq. (15)], giving the spectral weight of the pbz wave vector \vec{k}_0 in the SC eigenvector $|\vec{K} m\rangle$ as a quantity that can be obtained directly from a SC calculation alone.

*Current address: Faculty of Physics, University of Duisburg-Essen, Duisburg, Germany; voicu.popescu@uni-due.de

†alex.zunger@gmail.com

¹W. Shan, W. Walukiewicz, J. W. Ager III, E. E. Haller, J. F. Geisz, D. J. Friedman, J. M. Olson, and S. R. Kurtz, *Phys. Rev. Lett.* **82**, 1221 (1999).

²F. Masia, G. Pettinari, A. Polimeni, M. Felici, A. Miriametro, M. Capizzi, A. Lindsay, S. B. Healy, E. P. O'Reilly, A. Cristofoli *et al.*, *Phys. Rev. B* **73**, 073201 (2006).

³A. G. Thompson, M. Cardona, K. L. Shaklee, and J. C. Woolley, *Phys. Rev.* **146**, 601 (1966).

⁴J. Hwang, C. K. Shih, P. Pianetta, G. D. Kubiak, R. H. Stulen, L. R. Dawson, Y.-C. Pao, and J. J. S. Harris, *Appl. Phys. Lett.* **52**, 308 (1988).

⁵J. Endicott, A. Patanè, D. Maude, L. Eaves, M. Hopkinson, and G. Hill, *Phys. Rev. B* **72**, 041306 (2005).

⁶V. Popescu and A. Zunger, *Phys. Rev. Lett.* **104**, 236403 (2010).

- ⁷D. G. Thomas, J. J. Hopfield, and C. J. Frosch, *Phys. Rev. Lett.* **15**, 857 (1965).
- ⁸P. R. C. Kent and A. Zunger, *Phys. Rev. B* **64**, 115208 (2001).
- ⁹T. Mattila and A. Zunger, *Phys. Rev. B* **58**, 1367 (1998).
- ¹⁰H. Kamimura, in *Crystalline Semiconducting Materials and Devices*, edited by P. N. Butcher, N. H. March, and M. P. Tosi (Plenum Press, New York, 1986), pp. 305–354.
- ¹¹B. I. Shklovskii and A. L. Efros, *Electronic Properties of Doped Semiconductors*, vol. 45 of *Springer Series in Solid-State Sciences* (Springer Verlag, Berlin, 1984).
- ¹²P. W. Anderson, *Phys. Rev.* **109**, 1492 (1958).
- ¹³A. Luque and A. Martí, *Phys. Rev. Lett.* **78**, 5014 (1997).
- ¹⁴N. López, L. A. Reichertz, K. M. Yu, K. Campman, and W. Walukiewicz, *Phys. Rev. Lett.* **106**, 028701 (2011).
- ¹⁵A. Zunger, *J. Phys. C* **7**, 96 (1974).
- ¹⁶R. A. Evarestov, *Phys. Status Solidi B* **72**, 569 (1975).
- ¹⁷S. G. Louie and M. L. Cohen, *Phys. Rev. Lett.* **35**, 866 (1975).
- ¹⁸T. G. Dargam, R. B. Capaz, and B. Koiller, *Phys. Rev. B* **56**, 9625 (1997).
- ¹⁹L.-W. Wang, L. Bellaïche, S.-H. Wei, and A. Zunger, *Phys. Rev. Lett.* **80**, 4725 (1998).
- ²⁰T. B. Boykin, N. Kharche, G. Klimeck, and M. Korkusinski, *J. Phys.: Condens. Matter* **19**, 036203 (2007).
- ²¹W. Ku, T. Berlijn, and C.-C. Lee, *Phys. Rev. Lett.* **104**, 216401 (2010).
- ²²Y. Zhang and L.-W. Wang, *Phys. Rev. B* **83**, 165208 (2011).
- ²³A.-B. Chen and A. Sher, *Semiconductor Alloys: Physics and Materials Engineering* (Plenum Press, New York, 1995).
- ²⁴J. S. Faulkner and G. M. Stocks, *Phys. Rev. B* **21**, 3222 (1980).
- ²⁵H. Ebert, D. Ködderitzsch, and J. Minár, *Rep. Prog. Phys.* **74**, 096501 (2011).
- ²⁶Y. Zhang, A. Mascarenhas, and L.-W. Wang, *Phys. Rev. Lett.* **101**, 036403 (2008).
- ²⁷N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Saunders College Publishers, New York, 1976).
- ²⁸L. Bellaïche, S.-H. Wei, and A. Zunger, *Phys. Rev. B* **56**, 10233 (1997).
- ²⁹R. G. Dandrea and A. Zunger, *Phys. Rev. B* **43**, 8962 (1991).
- ³⁰V. Popescu and A. Zunger, *Phys. Rev. B* **84**, 125315 (2011).
- ³¹J. E. Bernard, S.-H. Wei, D. M. Wood, and A. Zunger, *Appl. Phys. Lett.* **52**, 311 (1988).
- ³²K. K. Bajaj, *Mater. Sci. Eng., R* **34**, 59 (2001).
- ³³P. N. Keating, *Phys. Rev.* **145**, 637 (1966).
- ³⁴R. M. Martin, *Phys. Rev. B* **1**, 4005 (1970).
- ³⁵K. Kim, P. R. C. Kent, A. Zunger, and C. B. Geller, *Phys. Rev. B* **66**, 045208 (2002).
- ³⁶A. J. Williamson, L. W. Wang, and A. Zunger, *Phys. Rev. B* **62**, 12963 (2000).
- ³⁷A. Zunger, *Quantum Theory of Real Materials* (Kluwer, Boston, 1996).
- ³⁸L.-W. Wang and A. Zunger, *J. Chem. Phys.* **100**, 2394 (1994).
- ³⁹S. V. Dudiy, P. R. C. Kent, and A. Zunger, *Phys. Rev. B* **70**, 161304 (2004).
- ⁴⁰S.-H. Wei and A. Zunger, *Appl. Phys. Lett.* **72**, 2011 (1998).
- ⁴¹T. Mattila, S.-H. Wei, and A. Zunger, *Phys. Rev. B* **60**, R11245 (1999).