Real-space description of semiconducting band gaps in substitutional systems

R. Magri and Alex Zunger

Solar Energy Research Institute, Golden, Colorado 80401

(Received 30 May 1991)

The goal of "band-gap engineering" in substitutional lattices is to identify atomic configurations that would give rise to a desired value of the band gap. Yet, current theoretical approaches to the problems, based largely on compilations of band structures for various latice configurations, have not yielded simple rules relating structural motifs to band gaps. We show that the band gap of substitutional AlAs/GaAs lattices can be usefully expanded in terms of a hierarchy of contributions from real-space "atomic figures" (pairs, triplets, quadruplets) detemined from first-principles band-structure calculations. Pair figures (up to fourth neighbors) and three-body figures are dominant. In analogy with similar cluster expansions of the total energy, this permits a systematic search among all lattice configurations for those having "special" band gaps. This approach enables the design of substitutional systems with certain band-gap properties by assembling atomic figures. As an illustration, we predict that the $[0\bar{1}2]$-oriented $(AlAs)_1/(GaAs)_4/(AlAs)_1/(GaAs)_2$ superlattice has the largest band gap among all $Al_{0.25}Ga_{0.75}As$ lattices with a maximum of ten cations per unit cell.

I. POSING THE PROBLEM

Substitutional A/B systems are the collection of crystal configurations that can be obtained by occupying the sites of a fixed (e.g., fcc) lattice by the atoms A and B. They contain ordered compounds, superlattices and substitutionally disordered alloys. Considerable research on substitutional semiconductor systems has focused on "band-gap engineering," i.e., on attempts to identify atomic configurations that lead to prescribed band gaps. Theoretical analysis of such problems has traditionally been based on band-structure techniques, whereby the dispersion of energy bands is mapped out in reciprocal space for a given real-space crystal configuration. In such approaches it has proved difficult to establish *intuitive rules* that relate band gaps to given motifs of atomic structure. Questions such as "what is the structure that, for a given composition, gives the largest direct band gap for AlAs/GaAs on a fcc lattice" have to be addressed, in principle, by calculating the band structures of a large number of configurations and selecting the one with the largest gap. What is clearly lacking here is a more direct connection between the band gap and the atomic configuration. There is, however, an established methodology relating the atomic configuration of a substitutional system to its *total* energy, i.e., cluster expansions in lattice models.[1]

Lattice (e.g., Ising) models of phase stability[1] exploit the fact that the many-electron ground-state energy surface of a substitutional system

$$E\{\hat{S}_i\} = \langle \Psi|\hat{H}|\Psi\rangle / \langle \Psi|\Psi\rangle \tag{1}$$

can be usefully parametrized in terms of a limited set of interatomic interaction energies J, i.e.,

$$E\{\hat{S}_i\} = J_0 + J_1 \sum_i S_i + \sum_{ij} J_{ij}\hat{S}_i\hat{S}_j + \sum_{ijk} J_{ijk}\hat{S}_i\hat{S}_j\hat{S}_k + \cdots . \tag{2}$$

Here, $\hat{S}_i$ is the pseudospin variable denoting in a binary A-B system whether site i is occupied by atom A ($\hat{S}_i = -1$) or B ($\hat{S}_i = +1$), and $\{\hat{S}_i\}$ is a particular occupation of the N sites by A and B atoms (a "configuration" σ). Establishment of a converged representation of $E(\sigma)$ in terms of a reasonably small number of interaction energies provides then a way for exploring the total energies of many more configurations[2] ($\le 2^N$ for a binary system) than is practical through solutions of (1) for all σ's. Furthermore, the renormalization of the explicit electronic degrees of freedom underlying the electronic Hamiltonian $\hat{H}$ into a digitized ("A on"/"B on") energy surface (2) permits the assessment of the importance of various atomic figures (e.g., ij pairs, ijk triplets) to the stability of various configurations.[3,4] These concepts were recently applied in the context of *first-principles* electronic-structure methods.[3-8] Applications to fcc lattices of binary transition metal[3-8] or semiconductor[4] systems showed that in most cases 5–10 interaction energies suffice to describe $E(\sigma)$ with useful precision, and that such a set of interaction energies can be used to search systematically the configurational space to identify σ's with, e.g., extremal values of the energy $E(\sigma)$.

In this paper we explore the possibility of using such a cluster expansion description for the *one-electron* orbital energies

$$\varepsilon_\alpha\{\hat{S}_i\} = \langle \phi_\alpha|\hat{H}|\phi_\alpha\rangle / \langle \phi_\alpha|\phi_\alpha\rangle \tag{3}$$

rather than the *total* energy [Eq. (1)]. More specifically, we will be interested in exploring whether the one-electron energy band gaps (i.e., differences between occupied and unoccupied orbital energies) can be usefully described in terms of contributions from a hierarchy of real-space "*figures*" of atoms:

$$E_g\{\hat{S}_i\} = p_0 + p_1 \sum_i S_i + \sum_{ij} p_{ij}\hat{S}_i\hat{S}_j + \sum_{ijk} p_{ijk}\hat{S}_i\hat{S}_j\hat{S}_k + \cdots . \tag{4}$$

As was proved by Sanchez, Ducastelle, and Gratias,[9] *any* lattice property can be *rigorously* expanded by the form (2) or (4), provided that the expansion in figures is carried to completeness (2^N terms). The practical usefulness of such expansions rests on the possibility of identifying a readily calculable and reasonably small set of ("building blocks") energies $\{p\}$ that span $E_g(\sigma)$ of arbitrary substitutional lattice configurations with useful precision (e.g., a few percent of E_g). If successful, this will provide insights into the way that band gaps reflect contributions from real-space atomic figures (rather than the conventional k-space picture) and enable the design of lattices with prescribed band gaps. This approach has a number of clear limitations. Since we use a *fixed lattice* description, we will insist that all lattice sites are occupied (so "broken bonds" are excluded) and that the Bravais symmetry is fixed so the coordination number is fixed too (although the calculations can be repeated for differently coordinated lattices). Consequently, we will not address here band-gap changes associated with (i) the formation of "dangling bonds" (as in certain amorphization or glass-forming processes), (ii) the creation of "antisite defects" [as in the randomization of Ga and As in zinc-blende GaAs (Ref. 10)], or (iii) from alotropic variations related to different coordination numbers[11] (e.g., zinc-blende versus rocksalt forms of GaAs,[12] the NaCl versus CsCl forms of Tl halides,[13] or zinc-blende versus the high-temperature forms of III-V semiconductors[14]). Even barring from our discussion such cases where the topology of the underlying lattice is altered in a fundamental way, it is still difficult to assess whether band gaps of substitutional systems depend primarily on short- or long-range order. On one hand, band-structure theory implies that gaps reflect by necessity the long-range order underlying the space group symmetry. Yet, inspection of the observed band gaps of fixed coordination number polytypes that have identical local environments up to the second neighbors [e.g., zinc-blende versus wurtzite, see Table I (Refs. 15–24)] reveals but small changes, suggesting perhaps that the short-range order has a dominant effect on the band gap of a given symmetry[25] (e.g., Γ_1). The crucial role of *local* topology in determining trends in band gaps has also been emphasized in the context of continued random network models of Si and Ge, where the existence of fivefold and sevenfold rings was correlated with increases in the band gaps.[26,27] Experimentally documented examples of variations in band gaps associated with *substitutional* modifications include the chalcopyrite–to–zinc-blende transition in $ZnSnP_2$, where $E_g = 1.22$ eV in the disordered zinc-blende phase and 1.46 eV in the ordered chalcopyrite structure[28] (1.64 eV according to Ref. 29). Analogous changes are observed in short-period isovalent semiconductor superlattices: the [001] $(AlAs)_1/(GaAs)_1$ superlattice has a low-temperature direct gap of 2.15 (Ref. 30) or 2.18 (Refs. 31 and 32) eV whereas when randomized (creating the $Al_{0.5}Ga_{0.5}As$ alloy), the direct gap rises to ~ 2.20 eV.[33] The 0.765-eV photoluminescence gap of the [001] superlattice $(InAs)_1/(GaAs)_1$ is increased by 35 meV upon substitutionally randomizing the system,[34] whereas the 1.85-eV room-temperature band gap of the [111] superlattice

$(GaP)_1/(InP)_1$ is increased by 85 (Ref. 35) or 140 meV (Ref. 36) as the system disorders. These changes are now well understood[32,37–40] in terms of *momentum-space* constructs. Depending on the superlattice orientation, different zinc-blende k points are folded into the superlattice Brillouin-zone center $\overline{\Gamma}$. For example, the zinc-blende $\Gamma + X$ states fold into $\overline{\Gamma}$ for [001] monolayer superlattices, whereas the zinc-blende states $\Gamma + L$ fold into $\overline{\Gamma}$ for the [111]-oriented superlattice. These folded states interact through the superlattice perturbation potential, resulting in a level repulsion. This effect overwhelms the opposite effect of quantum confinement thus leading to a reduction of the band gap of the superlattice relative to the random alloy. The reduction is largest for [111] superlattices since the unperturbed Γ and L energies are rather close. It is smaller in [001] superlattices, where the unperturbed Γ and X energies are farther apart. Our central challenge here is to depict such momentum-space events through real-space interactions between atoms.

II. THE DATA BASE

To explore the possibility of describing substitution-induced changes in band gaps through lattice models we first create an internally consistent data base of band structures for different substitutional crystal structures. We have selected the AlAs/GaAs system for a number of reasons. (i) The end-point constituents are very nearly lattice matched, so changes in relative composition leave the unit-cell volume unchanged. Hence all calculations can conveniently be performed at a constant molar volume. (ii) The direct band gaps of GaAs and AlAs differ considerably (by ~ 1.6 eV), so the overall changes

TABLE I. Measured direct band gaps (in eV) of the zinc-blende and wurtzite modifications of some II-VI semiconductors. These structures have identical local environments out to the second neighbor. The fact that their band gaps are similar (Refs. 15 and 16) appears then to suggest that, in some sense, the band gap is a "local" property.

	Zinc blende $\Gamma_{8v} \rightarrow \Gamma_{6c}$	Wurtzite $(\Gamma_{9v}, \Gamma_{7v} \rightarrow \Gamma_{7c})$
ZnS	3.74[a]	3.74[b] $(E\|c)$
		3.78 $(E\perp c)$
ZnSe	2.70[c]	2.795[d] $(E\perp c)$
	2.82[e]	2.874[e]
CdS	2.50[f]	2.501[g] (A)
		2.516 (B)
CdSe		1.751[h] (A)
		1.771 (B)

[a] $T = 295$ K, reflectance, Ref. 17.
[b] $T = 298$ K, reflectance, Ref. 18.
[c] $T = 300$ K, wavelength modulated reflectance, Ref. 17.
[d] $T = 300$ K, photoconductivity, Ref. 19.
[e] $T = 4.2$ K, exciton emission, Ref. 20.
[f] $T = 298$ K, reflectance, Ref. 21.
[g] $T = 300$ K, exciton absorption, Ref. 22.
[h] $T = 293$ K, exciton reflectance, Ref. 23.

of band gap with composition are sizable. (iii) The status of our experimental[30,31,33] and theoretical[32,37−40] understanding of the band gaps of disordered[33,41] and ordered[30−32,37−40] AlAs/GaAs systems is reasonably complete.

The crystal structures chosen reflect a range of compositions and ordering vectors. Denoting AlAs as "A" and GaAs as "B" we use the compositions A, A_3B, A_2B, AB, A_2B_2, AB_2, AB_3, and B. The layer orientations explored include [001], [110], [111], [201], and [113]. We have also included structures that are not superlattices, e.g., the AlGa$_3$As$_4$ (denoted L1) and Al$_3$GaAs$_4$ (L3) luzonite structures. Table II describes the various structures and assigns to them arbitrary names [Refs. 3(a) and 32 give pictures of many of these structures]. We have calculated the electronic band structure of these AlAs/GaAs systems using the first-principles nonlocal pseudopotential method[42] in the local-density formalism.[43] The scalar-relativistic pseudopotentials were generated by the method of Ref. 44. We use a plane-wave basis set with kinetic energy cutoff of 15 Ry and the exchange-correlation functional given by Perdew and Zunger.[45] Special care was exercised in sampling the Brillouin zone for different structures in a precisely equivalent manner:[32] we used for each structure a k point set equivalent to 29 zinc-blende k points[46] in the irreducible Brillouin zone. The theoretically calculated equilibrium lattice constants a_{eq} of AlAs and GaAs are very similar so we use for all structures the constant

atomic volume $a_{eq}^3 M/8$, where M is the number of atoms per cell and $a_{eq} = 10.619$ a.u. is the average calculated cubic lattice constant. Despite the well-known[45] underestimation of band gaps by the local-density approximation (LDA) our results are *internally* consistent to within a precision of better than 10 meV over the entire data base of structures. Because of the LDA error, we will judge the success of the cluster equation approach by the extent to which it reproduces the internally consistent LDA results, not experiment. It is possible, however, to approximately correct the LDA results by applying a rigid shift. As was noted previously,[32,33−40,47] the LDA calculated results for the direct gap at $\bar{\Gamma}$ are underestimated by a nearly constant value: the scalar relativistically calculated (experimental) values are 2.18 eV for AlAs (3.13 eV), 0.69 eV for GaAs (1.52 eV), 1.23 eV for the CA structure of GaAlAs$_2$ [2.15 eV (Refs. 30–32)], 1.38 eV for the Z2 structure Ga$_2$Al$_2$As$_4$ [∼2.19 eV (Refs. 30–32)]. (See Table II for the definition of the structures CA,Z2, etc.) An estimate of the corrected (including spin-orbit effect) band gap at $x = \frac{1}{2}$ can hence be obtained by adding ∼0.81 eV to the LDA results without spin orbit (this includes −0.11 eV for spin orbit and +0.92 eV for the direct LDA correction). This was discussed in some detail in Ref. 40.

Since we are addressing here *optical* transitions, we focus on the direct gap at $\bar{\Gamma}$ (the valence-band maximum always occurs in this system at the $\bar{\Gamma}$ point). To select the conduction band of zinc-blende Γ_{1c} symmetry we

TABLE II. Definition of the crystal structures used here for $A = $AlAs and $B = $GaAs. The structures in this table are described as A_pB_q superlattices with layer repeats (p,q) and orientation G. Structures that are *not* superlattices are pure A and pure B, and the AB_3 and A_3B (L1 and L3, respectively), both in the luzonite structure. Note that the A_1B_1 superlattice in the [001] direction (CA) is structurally identical to the monolayer superlattices in the [110] and [201] directions and that similar structural identities exist elsewhere (e.g., AB_2 and A_2B superlattices in the [110], [201], and [113] directions). The numbers in the table are the pseudopotential-calculated direct band gaps (in eV) at $\bar{\Gamma}$ at a unit-cell volume corresponding to the cubic lattice constant of 10.619 a.u. We have averaged over the small crystal-field splitting at the top of the valence band and considered the lowest-energy conduction band at $\bar{\Gamma}$ with a nonvanishing projection on a Γ_{1c} representation. The calculated band gaps for the structures not given in the table are 2.18 eV (A), 0.69 eV (B), 1.00 eV (L1), and 1.36 eV (L3).

Formula	Orientation [111]	[001]	[110]	[201]	[113]
AB	CP	CA	CA	CA	CP
	1.03	1.23	1.23	1.23	1.03
AB_2	α1	β1	γ1	γ1	γ1
	0.97	1.07	1.15	1.15	1.15
A_2B	α2	β2	γ2	γ2	γ2
	1.28	1.39	1.63	1.63	1.63
AB_3	V1	Z1	Y1	F1	W1
	0.88	0.99	1.01	1.04	0.92
A_2B_2	V2	Z2	Y2	CH	W2
	1.19	1.27	1.40	1.46	1.38
A_3B	V3	Z3	Y3	F3	W3
	1.37	1.60	1.67	1.69	1.52

have calculated the projection of a number of conduction-band states in each configuration onto a Γ_{1c} representation. We use the lowest-energy conduction band at $\overline{\Gamma}$ that has a finite Γ_{1c} projection. Note that in many cases this does not correspond to the absolute minimum of the conduction band at Γ, or to the conduction band with maximum Γ_{1c} projection. This produces, however, a consistent set of minimum-energy Γ_{1c}-like states appropriate to the onset of absorption or reflectance experiments. We have also averaged over the crystal-field splitting of the valence-band maximum (calculated as 42, 23, 0.6, and 16.6 meV for CA, CP, CH, and $Z2$, respectively; see Table II for the definitions of these structures).

The LDA calculated $\Gamma_{15v} \rightarrow \Gamma_{1c}$ one-electron band gaps are given in Table II and its caption for the 27 structures. The various symbols in Fig. 1 show these direct band gaps as a function of composition and superlattice

orientation. These exhibit (i) a substantial configurational dependence reflected in a spread in the values; (ii) some systematic trends, e.g., [111] structures have the smallest gaps while [201] structures have the largest gaps; mono-layer alternation (CP, CA) leads to smaller gaps than bilayer alternation ($V2$, $Z2$), etc.; (iii) the band gaps are significantly asymmetric with respect to the mid composition $x = \frac{1}{2}$; and (iv) all ordered structures except CH have smaller gaps than the linear concentration-weighted value $xE_g(A) + (1-x)E_g(B)$ [dashed line in Fig. 1(b)]. In what follows we wish to categorize such trends in terms of real-space lattice expansions.

III. CLUSTER EXPANSIONS

A binary A/B lattice with N sites can exist in 2^N different configurations σ corresponding to the various occupations of the N sites by the atoms A and B. We wish to expand a lattice property $P(\sigma)$ in a series of contributions p_f of "figures" f, defined as a cluster of atoms
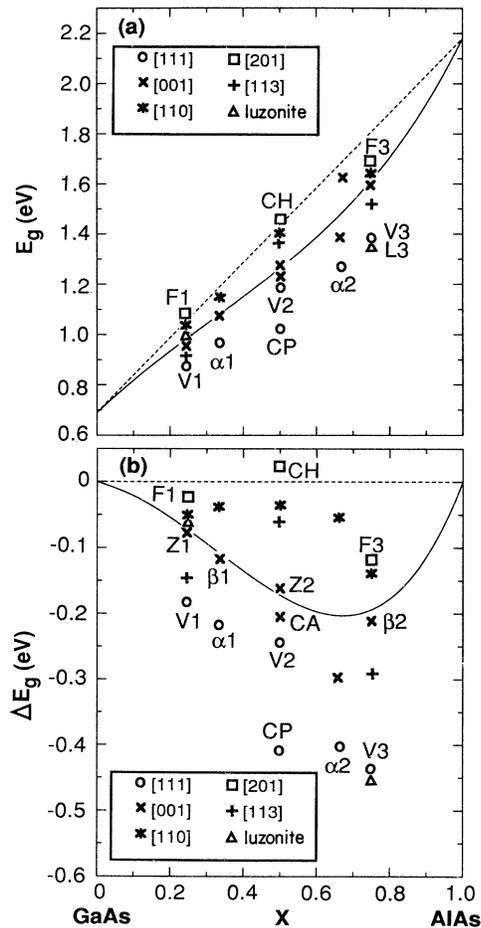

FIG. 1. Pseudopotential-calculated direct band gaps (symbols) of AlAs/GaAs superstructures with different layer orientations. See Table II for definitions of structural symbols. (a) Absolute values, (b) band gaps relative to the concentration-weighted values $xE_g(\text{GaAs}) + (1-x)E_g(\text{AlAs})$ shown as a dashed line. The solid line depicts the band gap of the random alloy as obtained from the cluster expansion using the full fit (Table V).
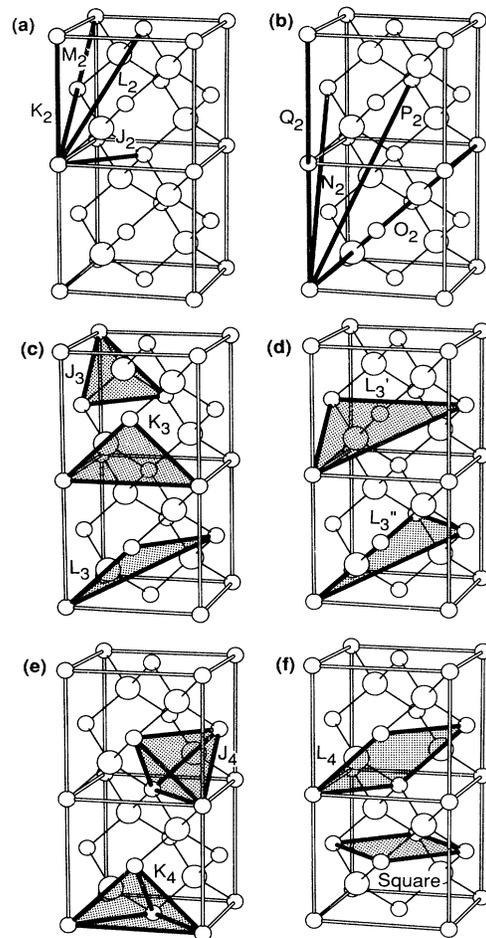

FIG. 2. Atomic figures used in the cluster expansion. See Table III for the definition of the figure coordinates, and Table IV for the nomenclature used to designate the various figures. (a) and (b): pair figures; (c) and (d): three-body figures; (e) and (f): four-body figures.

with k_f vertices (i.e., a selection of k_f out of N sites). Figure 2 depicts a number of such clusters in the fcc lattice and Table III defines their vertices. The expansion is defined with respect to an orthonormal set of coefficients, as follows. One defines the spin product

$$\Pi_f(\sigma) = \hat{S}_{1_f} \hat{S}_{2_f} \cdots \hat{S}_{k_f} , \tag{5}$$

for each of the 2^N figures in configuration σ. The set $\{\Pi_f(\sigma)\}$ is orthonormal,[9] [including the "empty figure" $f = (0,1)$ for which $\Pi_{0,1}(\sigma) = 1$] in that for two figures f and f' we have

$$\sum_\sigma \Pi_f(\sigma) \Pi_{f'}(\sigma) = 2^N \delta_{f,f'} . \tag{6}$$

Equation (6) shows that $2^{-N} \Pi_f(\sigma)$ is the transpose matrix of $\Pi_f(\sigma)$; multiplying the matrices in reverse order one obtains the completeness condition between the two configurations σ and σ',

$$\sum_f \Pi_f(\sigma) \Pi_f(\sigma') = 2^N \delta_{\sigma,\sigma'} . \tag{7}$$

One can hence rigorously expand any property $P(\sigma)$ of the lattice configuration σ in the orthonormal set of $\{\Pi_f(\sigma)\}$ as

$$P(\sigma) = \sum_{f=1}^{2^N} \Pi_f(\sigma) p_f , \tag{8}$$

where the configuration-independent contribution p_f of

figure f to the lattice property P is given from Eqs. (6)–(8) by

$$p_f = 2^{-N} \sum_{\sigma=1}^{2^N} \Pi_f(\sigma) P(\sigma) . \tag{9}$$

The series (8) may be reduced using symmetry: denoting by $\hat{R}$ one of the N_L operations of the space group of the lattice (not the space group of a particular configuration σ), we have

$$P(\hat{R}\sigma) = P(\sigma) \tag{10}$$

and

$$\Pi_{\hat{R}f}(\hat{R}\sigma) = \Pi_f(\sigma) , \tag{11}$$

hence, Eq. (9) gives

$$p_{\hat{R}f} = p_f , \tag{12}$$

so all ND_F symmetry-related figures contribute equally to $P(\sigma)$. This fact can be used to reduce the sum in Eq. (8) to just the symmetry-inequivalent figures F,

$$P(\sigma) = N \sum_F \overline{\Pi}_F(\sigma) D_F p_F , \tag{13}$$

where the "lattice-averaged spin product" (denoted by an overbar) of the prototype figure F in configuration σ is

$$\overline{\Pi}_F(\sigma) = \frac{1}{N_L} \sum_{\hat{R}}^{N_L} \Pi_{\hat{R}F}(\sigma) , \tag{14}$$

TABLE III. Definition of the vertices of the figures (Fig. 2) considered here for fcc lattices.

Figure	Site positions (in units of $a = 2$)		
(0,1)	Empty		
(1,1)	(0,0,0)		
Pairs			
$(2,1) = J_2$	(0,0,0) (1,1,0)		
$(2,2) = K_2$	(0,0,0) (2,0,0)		
$(2,3) = L_2$	(0,0,0) (2,1,1)		
$(2,4) = M_2$	(0,0,0) (2,2,0)		
$(2,5) = N_2$	(0,0,0) (3,1,0)		
$(2,6) = O_2$	(0,0,0) (2,2,2)		
$(2,6) = P_2$	(0,0,0) (3,2,1)		
$(2,8) = Q_2$	(0,0,0) (0,4,0)		
Three body			
$(1,1,1) = J_3$	(0,0,0) (1,1,0) (1,0,1)	
$(2,1,1) = K_3$	(0,0,0) (1,1,0) (2,0,0)	
$(3,1,1) = L_3$	(0,0,0) (1,1,0) (2,1,1)	
$(3,2,1) = L_{3'}$	(0,0,0) (1,1,0) (1,1,2)	
$(3,3,1) = L_{3''}$	(0,0,0) (1,1,0) (−1,2,1)	
Four body			
$(4,1) = J_4$	(0,0,0) (1,1,0) (1,0,1) (0,1,1)
$(4,2) = K_4$	(0,0,0) (1,1,0) (1,0,1) (2,0,0)
$(4,3) = L_4$	(0,0,0) (1,1,0) (1,0,1) (2,1,1)
Square	(0,0,0) (1,1,0) (−1,1,0) (0,2,0)

and N_L is the number of operations $\hat{R}$ in the lattice space group (e.g., $N_L = 48N$ for fcc lattices). The set $\{\overline{\Pi}_F(\sigma)\}$ characterizes unequivocally the structure of configuration σ. The cluster expansion of Eq. (8) defines a multisite Ising Hamiltonian[1] which includes "interactions" $p_F \equiv p_{k,m}$ between k sites separated by up to the mth-neighbor distance (the choice $k = 2$, $m = 1$ corresponds to the classic nearest-neighbor pair interaction case).

It is useful to expand the configurational property $P(\sigma)$ with respect to some reference configuration. One possibility is to expand it relative to the property P of equivalent amounts of pure A and pure B lattices. The *excess* property ΔP for the $A_{1-x}B_x$ system (with composition x) with respect to equivalent amounts of A and B is

$$\Delta P(\sigma) = P(\sigma) - [(1-x)P(A) + xP(B)] . \tag{15}$$

The cluster expansion for ΔP is then

$$\Delta P(\sigma) = \sum_{m>0} \sum_{k>1} [\overline{\Pi}_{k,m}(\sigma) - \eta] D_{k,m} p_{k,m} , \tag{16}$$

where $\eta = 1$ for k even and $\eta = (2x - 1)$ for k odd.

For a perfectly random (R) alloy, one needs to consider an ensemble average over the 2^N configurations. This gives

$$\langle \overline{\Pi}_{k,m} \rangle_R = (2x - 1)^k . \tag{17}$$

Phenomenologically, the band gap of the random $A_{1-x}B_x$ alloy has traditionally been expressed as[33,41]

$$E_g^{(R)}(x) \equiv [(1-x)E_g(A)+xE_g(B)]-bx(1-x) , \quad (18)$$

where b is the "bowing" coefficient. Using Eqs. (15)–(18) we can then express b as a series of interactions:

$$b(x)=b_{2\text{-body}}+b_{3\text{-body}}+b_{4\text{-body}}+ \cdots , \quad (19)$$

where

$$b_{2\text{-body}}= \sum_{m>0} 4D_{2,m}p_{2,m} ,$$

$$b_{3\text{-body}}=(2x-1) \sum_{m>0} 4D_{3,m}p_{3,m} , \quad (20)$$

$$b_{4\text{-body}}=[(2x-1)^2+1] \sum_{m>0} 4D_{4,m}p_{4,m} ,$$

and so on. Alternatively, one can write

$$b(x)=b_0+(2x-1)b_1+(2x-1)^2b_2+ \cdots \quad (21)$$

where

$$b_0= \sum_{m>0} (4D_{2,m}p_{2,m}+4D_{4,m}p_{4,m})+ \cdots ,$$

$$b_1= \sum_{m>0} 4D_{3,m}p_{3,m}+ \cdots , \quad (22)$$

$$b_2= \sum_{m>0} 4D_{4,m}p_{4,m}+ \cdots .$$

This shows that (i) nonzero bowing is a statement of existence of many-body (including pair) interactions since $p_{0,1}$ and $p_{1,1}$ describe the purely linear composition dependence, i.e., the first bracketed term in Eq. (18). (ii) The composition-independent piece b_0 of the bowing parameter (assumed to be the only piece in most phenomenological treatments) reflects even-body interactions; composition dependence of b then arises from three- and more-body effects. (iii) Odd-body corrections to optical bowing vanish at $x=\frac{1}{2}$.

While the complete cluster expansions of Eqs. (8) or (13) are formally exact, they merely replace a direct calculation of 2^N values of $P(\sigma)$ by an equivalent number of calculations of the elementary contributions $\{p_f\}$. The utility of these expansions rests, however, on the possibility of identifying a hierarchy of a small number $(\ll 2^N)$ of figures whose contributions $p_{k,m}$ to the lattice property P dominates those of the remaining figures. Note the limitation here: in general, one could not predict $P(\sigma)$ for configurations σ whose unit-cell size exceeds the dimension of the largest figure included in the expansion (unless $p_{k,m}$ is negligible for larger figures). To the extent that the basic cluster expansion of Eq. (13) converges regularly and rapidly with respect to the figures (a point to examine below), one can use any sufficiently large set of $\{P(\sigma)\}$ in Eq. (9) to evaluate the effective cluster properties $\{p_F\}$. Conversely, nonunique values of $\{p_F\}$ obtained from two different sets of configurations $\{\sigma\}$ and $\{\sigma'\}$ of comparable sizes testify to the importance of interactions beyond the truncation limit used. This suggests that one can (i) establish a trial maximum figure $F_{\max}$ to be retained in the cluster expansion of Eq. (13), (ii) select a computationally convenient set of configuration $\{\sigma\}$ (e.g.,

periodic structures) from which p_F for $F \le F_{\max}$ can be obtained, and (iii) examine convergence by using $\{p_F\}$ to predict the property $P(\sigma')$ for other structures $\{\sigma'\} \neq \{\sigma\}$; if this fails, $F_{\max}$ is increased until transferability is established.

Generalizing the Connolly and Williams[5] approach we specialize the expansion of Eq. (13) to a set of N_s periodic structures $\{\sigma\}=\{s\}$ for which (i) $P(\sigma)$ can be readily calculated (e.g., from band theory), and (ii) $\{\overline{\Pi}_F(s)\}$ are known trivially. One then obtains the effective cluster properties $\{p_F\}$ by minimizing the weighted variance,

$$\sum_{s=1}^{N_s} \omega_s \left[P(s)- \sum_F^{N_F} \overline{\Pi}_F(s)D_F p_F \right]^2 = \text{minimum} , \quad (23)$$

with the weights

$$\omega_s =48N_c(s)/N_G(s) . \quad (24)$$

Here $N_c(s)$ and $N_G(s)$ are the number of atoms per unit cell and the number of point group operations for the structure s, respectively. Our basic strategy will be to use the band gaps $\{P(s)\}$ calculated from band-structure theory for a set of ordered structures $\{s\}$ (Table II) and obtain $\{p_F\}$ from Eqs. (23) and (24). We will then examine convergence and transferability. A converged set of
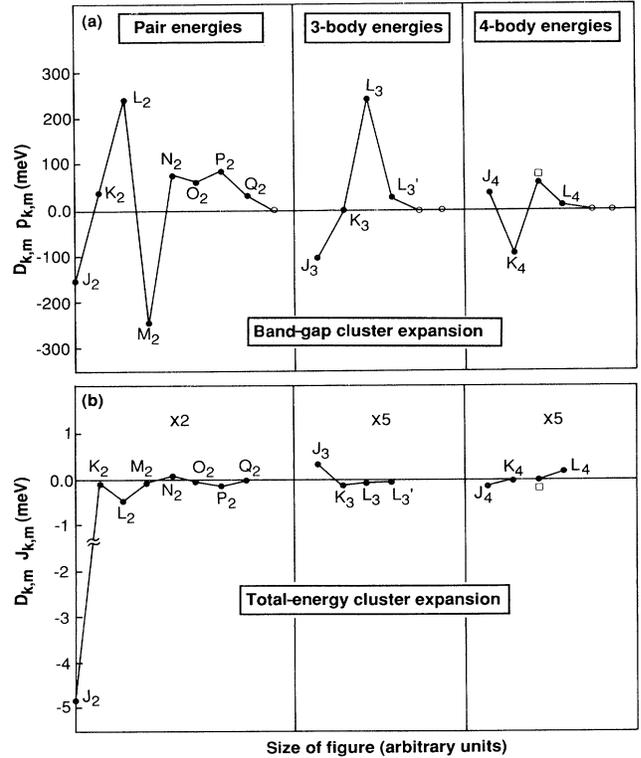

FIG. 3. Product of the degeneracy factor $D_{k,m}$ and the interaction coefficient $p_{k,m}$. (a) Cluster expansion of the direct band gap of AlAs/GaAs systems ($D_{k,m}$ and $p_{k,m}$ of Table VI). (b) Cluster expansion of the excess total energy of AlAs/GaAs systems ($D_{k,m}$ and $J_{k,m}$ of Table VI).

$\{p_F\}$ can then be used to predict band gaps of arbitrary structures.

IV. INTERACTION CONSTANTS FOR THE AlAs/GaAs SYSTEMS

The pseudopotential calculated band gaps $E_g(s)$ of our 27 structures (Table II) were used in Eq. (23) to obtain the interaction energies $\{p_{k,m}\}$. Table IV gives the correlation functions $\overline{\Pi}_{k,m}(s)$ for these and other structures as well as the weights ω_s. Since there is no prior experience as to the rate of convergence of the band-gap cluster expansion,[48] we started with a rather large set of 18 figures, shown in Fig. 2. In addition to the normalization constant $p_{0,1}$ and the site-only $p_{1,1}$ figures we use (i) eight pair $(k=2)$ interactions [Figs. 2(a) and 2(b)], (ii) four three-body $(k=3)$ figures [Figs. 2(e) and 2(d)] and (iii) four four-body $(k=4)$ figures [Figs. 2(c) and 2(f)]. The resulting fit of 18 interactions to 27 band gaps is given in Table V under the heading "full fit." Table VI gives the interaction energies $\{p_{k,m}\}$ along with the values obtained by fitting the band gap obtained in a virtual crystal approximation (VCA). The latter is nearly linear with composition, so all of the many-body terms are negligible. The product of the degeneracy factor $D_{k,m}$ and the interaction energy $p_{k,m}$ is shown graphically in Fig. 3.

TABLE IV. Correlation functions $\overline{\Pi}_{k,m}(s)$ [Eq. (14)] for figures (k,m) (shown in Fig. 2 and Table III) in structures s (defined in Table II). $D_{k,m}$ is the degeneracy factor. To obtain the results for AB_2C_3 from those of A_2BC_3, switch the sign of the odd-body $(k=1,3)$ figures. The same applies to $A\rightarrow B$ and $A_3BC_4\rightarrow AB_3C_4$ structures. For the (R) random alloys, $q=2x-1$.

Composition		AC	ABC₂		A₂BC₃			A₂B₂C₄							A₃BC₄					A₄B₄C₈		R
Name		A	CA	CP	α_1	β_1	γ_1	CH	V_2	W_2	Y_2	Z_2	F_1	L_1	V_1	W_1	X_1	Y_1	Z_1	D_4	SQS-8	
Weight ω_s		1	6	8	12	9	18	12	16	48	24	12	12	4	16	48	24	24	12	8	96	
Figures	$D_{k,m}$																					
$(0,1)=J_0$	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
$(1,1)=J_1$	1	-1	0	0	$-\frac13$	$-\frac13$	$-\frac13$	0	0	0	0	0	$-\frac12$	$-\frac12$	$-\frac12$	$-\frac12$	$-\frac12$	$-\frac12$	$-\frac12$	0	0	q
Pairs:																						
$(2,1)=J_2$	6	1	$-\frac13$	0	$\frac13$	$\frac19$	$-\frac19$	$-\frac13$	$\frac12$	$-\frac16$	0	$\frac13$	0	0	$\frac12$	$\frac16$	$\frac16$	$\frac16$	$\frac13$	0	0	q^2
$(2,2)=K_2$	3	1	1	-1	$-\frac13$	$\frac59$	$\frac19$	$\frac13$	0	0	$-\frac13$	$\frac13$	$\frac23$	1	0	0	0	$\frac13$	$\frac23$	-1	0	q^2
$(2,3)=L_2$	12	1	$-\frac13$	0	0	$-\frac13$	$\frac13$	$\frac13$	0	0	0	$-\frac13$	$\frac13$	0	$\frac14$	$\frac14$	$\frac16$	$\frac16$	0	0	$\frac1{24}$	q^2
$(2,4)=M_2$	6	1	1	1	$\frac13$	$\frac19$	$-\frac19$	$-\frac13$	0	0	$-\frac13$	$-\frac13$	$\frac13$	1	$\frac12$	$\frac12$	1	$\frac13$	$\frac13$	1	$-\frac1{12}$	q^2
$(2,5)=N_2$	12	1	$-\frac13$	0	$-\frac13$	$\frac59$	$\frac19$	$-\frac13$	$-\frac12$	$\frac16$	0	$\frac13$	0	0	0	$\frac13$	$\frac16$	$\frac16$	$\frac13$	0	$\frac1{12}$	q^2
$(2,6)=O_2$	4	1	1	-1	0	$-\frac13$	$\frac13$	-1	0	0	1	-1	0	1	0	0	0	1	0	-1	0	q^2
$(2,7)=P_2$	24	1	$-\frac13$	0	$\frac13$	$\frac19$	$-\frac19$	$\frac13$	0	0	0	$-\frac13$	$\frac13$	0	$\frac14$	$\frac14$	$\frac16$	$\frac16$	0	0	$-\frac18$	q^2
$(2,8)=Q_2$	3	1	1	1	$-\frac13$	$\frac59$	$\frac19$	1	-1	-1	1	1	1	1	0	0	1	1	1	1	$-\frac12$	q^2
$(2,9)=R_2$	12	1	$-\frac13$	0	0	$-\frac13$	$\frac13$	$-\frac13$	$-\frac12$	$\frac16$	0	$\frac13$	0	0	0	$\frac13$	$\frac16$	$\frac16$	$\frac13$	0	$\frac1{12}$	q^2
$(2,9)=R_2'$	6	1	$-\frac13$	0	1	1	1	$-\frac13$	$\frac12$	$-\frac16$	0	$\frac13$	0	0	$\frac12$	$\frac16$	$\frac16$	$\frac16$	$\frac13$	0	0	q^2
$(2,10)=S_2$	12	1	1	-1	$\frac13$	$\frac19$	$-\frac19$	$\frac13$	0	0	$-\frac13$	$\frac13$	$\frac23$	1	0	0	0	$\frac13$	$\frac23$	-1	0	q^2
Three body:																						
$(1,1,1)=J_3$	8	-1	0	0	$-\frac13$	$-\frac13$	$\frac13$	0	0	0	0	0	$\frac12$	$\frac12$	$-\frac12$	0	0	0	$-\frac12$	0	$-\frac14$	q^3
$(2,1,1)=K_3$	12	-1	0	0	$-\frac13$	$\frac19$	$\frac19$	0	0	0	0	0	$-\frac16$	$-\frac12$	$-\frac12$	$\frac16$	$\frac16$	$-\frac16$	$-\frac16$	0	0	q^3
$(3,1,1)=L_3$	24	-1	0	0	0	$\frac19$	$-\frac19$	0	0	0	0	0	$\frac16$	$\frac12$	$-\frac14$	$-\frac1{12}$	0	0	$-\frac16$	0	$\frac18$	q^3
$(3,2,1)=L_3'$	24	-1	0	0	0	$-\frac13$	$\frac19$	0	0	0	0	0	$-\frac12$	$-\frac12$	$-\frac14$	$\frac1{12}$	$\frac16$	$\frac16$	$-\frac12$	0	$-\frac1{24}$	q^3
$(3,3,1)=L_3''$	24	-1	0	0	0	$\frac59$	$-\frac19$	0	0	0	0	0	$-\frac16$	$\frac12$	$-\frac12$	0	0	0	$\frac16$	0	$\frac1{12}$	q^3
Four body:																						
$(4,1)=J_4$	2	1	1	-1	$-\frac13$	1	$-\frac13$	1	0	0	-1	1	-1	-1	0	0	0	0	1	1	0	q^4
$(4,2)=K_4$	12	1	$-\frac13$	0	$\frac13$	$-\frac13$	$-\frac19$	$\frac13$	$\frac12$	$-\frac16$	0	$-\frac13$	$-\frac13$	0	$\frac12$	$-\frac16$	$-\frac16$	$\frac16$	0	0	$-\frac16$	q^4
$(4,3)=L_4$	12	1	1	1	$\frac23$	$\frac59$	$\frac19$	$\frac13$	$\frac12$	$\frac12$	$\frac13$	$\frac13$	$-\frac23$	-1	$\frac34$	$\frac1{12}$	0	0	$\frac23$	-1	$\frac5{24}$	q^4
Square	3	1	1	1	1	$\frac19$	$\frac19$	$-\frac13$	1	1	$-\frac13$	$-\frac13$	$\frac13$	1	1	$-\frac13$	$-\frac13$	$\frac13$	$\frac12$	1	$\frac13$	q^4

From these results, we conclude the following.

(i) The cluster expansion for the direct band gap [Fig. 3(a)] converges considerably slower than that for the total energy[49] [Eq. (2) and Fig. 3(b)] in that the latter is dominated by the nearest-neighbor pair energy $J_{2,2}$ while the former requires a considerably larger number of higher terms.

(ii) Nevertheless, the cluster expansion of the band gap is reasonably accurate in that the standard deviation for the fit (Table V) is comparable to the underlying precision of the pseudopotential calculation.

(iii) The cluster energies of Fig. 3(a) exhibit an overall decrease with size, in that the contributions from figures with larger interatomic separations are rather small.

V. CONVERGENCE AND TRANSFERABILITY

We have conducted four tests to examine the convergence of the cluster expansion.

First, we have recalculated $\{p_{k,m}\}$ from Eq. (23), using as input only 10 out of 27 structures; the resulting interaction energies were then used to predict the band gaps of the remaining 17 ordered structures *not used in the fit*. The last two columns of Table V show that the prediction error (relative to the "exact" pseudopotential values) has increased only to 0.04 eV compared to 0.03 eV in the complete fit; both deviations are close to the relative uncertainty in the pseudopotential calculations of the band gaps. Note that while both fits predict very similar band gaps, the nonunique values of $\{p_{k,m}\}$ can differ.

As a second test, we have calculated the band gap of the *random* $Al_{1-x}Ga_xAs$ alloy from Eqs. (13) and (17), using different sets of ordered structures for extracting $\{p_{k,m}\}$. Table VII shows that the random alloy gap is predicted reasonably robustly starting with different choices of ordered structures. In general, the band gaps of disordered alloys can be predicted with greater accuracy than those of ordered structures since the former

TABLE V: Pseudopotential-calculated and cluster expanded LDA band gaps using N_s structures and N_f figures. The "full fit" was carried out for $N_s = 27$ structures and the $N_f = 18$ interactions shown in Table VI. The "partial fit" was carried out for $N_s = 10$ structures (denoted here by an asterisk) and $N_f = 9$ interactions. These include J_0 and J_1, the first five pair interactions J_2, K_2, L_2, M_2, and N_2 as well as the three-body terms J_3 and L_3. "Difference" refers to the cluster expansion value minus the pseudopotential-calculated band gap. Values in the last two columns not indicated by an asterisk denote predicted quantities (i.e., not included in the fit). The last two lines give the average error (AE) over the N_s structures used in the fit ("AE of fit") and over the 27 structures ("AE of predictions").

Structures	Calculated (eV)	Full fit $N_s = 27$; $N_f = 18$		Partial fit $N_s = 10$; $N_f = 9$	
		Fitted (eV)	Difference (eV)	Fitted (eV)	Difference (eV)
GaAs	0.69	0.69	0.00	0.69*	0.00*
AlAs	2.18	2.18	0.00	2.18*	0.00*
CA	1.23	1.22	−0.01	1.17*	−0.06*
CP	1.03	0.99	−0.04	1.03*	0.00*
$\alpha1$	0.97	0.99	+0.02	0.97*	0.00*
$\alpha2$	1.28	1.30	+0.02	1.28*	0.00*
$\beta1$	1.07	1.14	+0.07	1.15	+0.08
$\beta2$	1.39	1.43	+0.04	1.40	+0.01
$\gamma1$	1.15	1.17	+0.02	1.17	+0.02
$\gamma2$	1.63	1.68	+0.05	1.66	+0.03
CH	1.46	1.45	−0.01	1.46*	0.00*
$V2$	1.19	1.14	−0.05	1.17	−0.02
$W2$	1.38	1.36	−0.02	1.32	−0.06
$Y2$	1.40	1.38	−0.02	1.30	−0.10
$Z2$	1.27	1.24	−0.03	1.27*	0.00*
$F1$	1.04	1.05	+0.01	1.11	+0.07
$F3$	1.69	1.67	−0.02	1.66	−0.03
$L1$	1.00	0.95	−0.05	1.06*	+0.06*
$L3$	1.36	1.40	+0.04	1.42*	+0.06*
$V1$	0.88	0.87	−0.01	0.90	+0.02
$V3$	1.37	1.48	+0.11	1.51	+0.14
$W1$	0.92	0.90	+0.02	0.90	+0.06
$W3$	1.52	1.53	+0.01	1.57	+0.05
$Y1$	1.01	1.06	+0.05	1.02	+0.01
$Y3$	1.67	1.63	−0.04	1.58	−0.09
$Z1$	0.99	0.99	−0.00	1.01*	+0.02
$Z3$	1.60	1.57	−0.03	1.57*	−0.03
AE of fit			+0.03		+0.02
AE of predictions			+0.03		+0.04

TABLE VI. Band-gap expansion coefficients $p_{k,m}$ and the analogous expansion coefficients $J_{k,m}$ for the *total* energy [Eq. (2)]. $p_{k,m}$ (VCA) gives the cluster coefficients for expansion of the virtual crystal approximation to the band gap.

Figure	$D_{k,m}$	$p_{k,m}$ (meV)	$P_{k,m}$ (VCA)	$J_{k,m}$
$(0,1)=J_0$	1	1262.86	1435.7	5.2638
$(1,1)=J_1$	1	584.06	740.9	-0.0059
Pairs				
$(2,1)=J_2$	6	-25.68	0.00	-0.8054
$(2,2)=K_2$	3	13.05	0.00	-0.0156
$(2,3)=L_2$	12	20.68	0.00	-0.0018
$(2,4)=M_2$	6	-41.44	0.00	-0.0044
$(2,5)=N_2$	12	6.40	0.00	0.0003
$(2,6)=O_2$	4	16.80	0.00	-0.0085
$(2,7)=P_2$	24	3.58	0.00	-0.0024
$(2,8)=Q_2$	3	11.05	0.00	-0.0027
Three body				
$(1,1,1)=J_3$	8	-12.92	0.00	0.0083
$(2,1,1)=K_3$	12	-0.90	0.00	-0.0025
$(3,1,1)=L_3$	24	10.36	0.00	-0.0008
$(3,2,1)=L_{3'}$	24	1.10	0.00	-0.0004
Four body				
$(4,1)=J_4$	2	17.80	0.00	-0.0106
$(4,2)=K_4$	12	-7.79	0.00	-0.0005
Square	3	22.80	0.00	-0.0004
$(4,3)=L_4$	12	1.17	0.00	0.0035

represent an average over ordered structures.

A third way to examine the cluster expansion involves the direct calculation (not a fit) of some $p_{k,m}$ values. Table IV shows that the structure $A_4B_4C_8$ denoted D_4 (Refs. 50–52) has identical pair ($k=2$) and three-body ($k=3$) correlation functions $\overline{\Pi}_{k,m}$ to that of the CP structure, for all of the considered orders of m; these structures differ only in some of the four-body terms. Using Eq. (13) and the $\overline{\Pi}_{k,m}$ values of Table IV we then have from the cluster expansion (CE)

$$E_g^{\text{CE}}(D_4)-E_g^{\text{CE}}(CP)=(4J_4-24L_4)\ \cdots \quad (25)$$

where $\cdots$ represents all higher many-body terms. The pseudopotential (PS) calculation gives

$$E_g^{(\text{PS})}(D_4)-E_g^{(\text{PS})}(CP)=-0.02\ \text{eV}\ . \quad (26)$$

Comparison of (25) with (26), using the values of J_4 and L_4 from Table VI then shows that all higher-order terms equal 0.06 eV, a reasonably small correction.

A fourth test consisted of using the cluster expansion to predict the band gap of a complex, low symmetry structure. We selected the SQS-8 structures (defined in Ref. 48) consisting of an $A_2B_3A_2B_1$ superlattice along the [113] direction, or the transposed $B_2A_3B_2A_1$ structure. The average of the band gaps of these two structures, as obtained by direct pseudopotential calculation, is 1.38 ± 0.03 eV. The cluster expansion using 27 ordered structures as a basis gave 1.33 eV, showing again reasonable convergence.

Based on these tests we conclude that our cluster expansion can predict the band gaps of arbitrary substitutional AlAs/GaAs structures (whose unit cells are contained within our largest figures) to within better than 0.1

TABLE VII. Predicted LDA band gap of the random $Al_{0.5}Ga_{0.5}As$ alloy using different sets of N_s ordered structures.

N_s	Structures used in expansion	Interaction terms	Alloy gap at $x=\frac{1}{2}$ (eV)
5	$A,B,CA,L1,L3$	J_0,J_1,J_2,J_3,J_4	1.231
6	$A,B,CA,CP,CH,Z2$	J_0,J_1,J_2,K_2,L_2,M_2	1.281
10	$A,B,CA,CH,CP,Z2$ $\alpha1,\alpha2,L1,L3$	J_0,J_1,J_2,K_2,L_2 M_2,N_2,J_3,L_3	1.281
27	27 structures (Table IV)	J_0,J_1,J_2,K_2,L_2 M_2,N_2,O_2,J_3 K_3,L_3,J_4,K_4	1.301

eV, i.e., just a bit above the precision of the underlying pseudopotential calculation.

VI. APPLICATIONS
OF THE CLUSTER EXPANSION

Having established the practical convergence and transferability of the band-gap cluster expansion, we next apply it to predict band gaps for structures that are too complex to be calculated via Eq. (3).

A. Predicting the direct band gap
of the $Al_{1-x}Ga_xAs$ random alloy

The solid lines in Fig. 1 show the predicted band gap of perfectly random $Al_{1-x}Ga_xAs$ solid solutions, using Eqs. (15)–(17) and the interaction energies of Table VI. We see that the bowing of $\langle E_g(x)\rangle_R$ is nonparabolic. Using Eqs. (19) and (20) we find

$$b_{2\text{-body}} = 0.59 \text{ eV} ,$$

$$b_{3\text{-body}} = 0.64(2x-1) \text{ eV} , \tag{27}$$

$$b_{4\text{-body}} = 0.098[(2x-1)^2 + 1] \text{ eV} ,$$

i.e., two- and three-body interactions are dominant. The comparison of these values to experiment is somewhat clouded by the LDA error on the side of theory, and by the fact that alloy data is usually displayed as an *average* over certain composition ranges, on the side of experiment. Recent careful experimental studies[41] have revealed, however, an unusual nonparabolic behavior of $\langle E_g(x)\rangle_R$, with a small and constant bowing in the GaAs-rich range and a significantly larger bowing at higher Al concentrations, in qualitative agreement with our results (Fig. 1). Recent tight-binding model calculations have shown small bowing[52(a)] in the GaAs-rich region; more refined tight-binding calculations[52(b)] revealed a strong *negative* bowing in the GaAs-rich alloy that is not seen in our calculation and in most recent experimental data.[41] Calculations based on the coherent potential approximation[53] show likewise a smaller bowing in the Ga-rich alloy ($b\sim0.1$ eV) than in the Al-rich alloy ($b\sim0.25$), while VCA calculations[54] (see also Table VI) reveal essentially no bowing in this system.

Figure 1(a) depicts the absolute values of the calculated band gaps, whereas Fig. 1(b) displays the values of $\Delta E_g(x)$, after subtracting the linear weighted average of the band gap of the constituents $xE_g(A)+(1-x)E_g(B)$. This highlights the fact that $\Delta E_g(x)$ is significantly asymmetric with respect to the midcomposition $x=\frac{1}{2}$ value, in sharp contrast with the highly symmetric excess *total* energy[49] $\Delta E_{tot}(x)$. Our foregoing analysis [see Eq. (27)] shows that this reflects significant odd-body interactions controlling the band gap [Fig. 3(a)] but not the total energy [Fig. 3(b)].

B. Effects of short-range order on the band gap
of disordered $Al_{1-x}Ga_xAs$

Equation (17) applies to substitutional alloys that are *perfectly* random, hence, so do our alloy results of Eqs. (19)–(22), and Fig. 1. Fu, Chao, and Osorio[55] and Has-

bun, Singh, and Roth[56] pointed out that short-range order (SRO) could affect the alloy band gap. While it was expected qualitatively that the $Al_{1-x}Ga_xAs$ alloy will tend to phase separate at low temperatures (hence the SRO will consist of a greater clustering tendency than that granted by random statistics), the precise form of SRO was not known. Fu, Chao, and Osorio[55] hence assumed a simple form of SRO consistent with clustering tendencies, finding that it led to significant *upward bowing* ($b<0$) of the direct band gap in $Al_{1-x}Ga_xAs$.

The effects of SRO on the alloy band gap can be described within our formalism, using accurately calculated values of the SRO. Wei, Ferreira, and Zunger[4(b)] performed a cluster expansion of the *total* energy [Eq. (1)] of AlAs/GaAs and used the resulting interaction energies J to calculate (via the cluster variation method) the temperature-composition phase diagram. They also obtained the excess clustering probabilities

$$\Delta Q_n(x,T) = Q_n(x,T) - Q_n^{(R)}(x,\infty) , \tag{28}$$

where $Q_n(x,T)$ is the probability of finding at (x,T) the As-centered clusters $Al_{4-n}Ga_n$ ($0 \leq n \leq 4$) and $Q_n^{(R)}(x,\infty)$ is the random probability (at $x=\frac{1}{2}$, this is $\frac{1}{16}$, $\frac{4}{16}$, $\frac{6}{16}$, $\frac{4}{16}$, and $\frac{1}{16}$ for $n=0$, 1, 2, 3, and 4, respectively). Figure 8 of Ref. 4(b) depicts that $\Delta Q_n(x,T)$, showing that at finite temperatures the alloy is enriched by the $n=0$ (Al_4) and $n=4$ (Ga_4) clusters over what random statistics would grant, while the *mixed* Al_3Ga, Al_2Ga_2, and $AlGa_3$ clusters are deficient. We next describe how such thermodynamic results could be incorporated into our cluster expansion for the band gaps, giving the effect of SRO on optical properties.

Our Eq. (13) can be reformulated as a sum over a set of ordered configurations $\{s\}$ rather than over figures $\{F\}$ by noting that the interaction energies p_F are

$$p_F = \frac{1}{ND_F} \sum_s [\overline{\Pi}_F(s)]^{-1} E_g(s) \tag{29}$$

and $E_g(s)$ is the band gap of structure s. Inserting Eq. (29) into Eq. (13) and taking the configurational average (denoted by angular brackets) appropriate to the disordered phase, one gets

$$\langle E_g \rangle_R = \sum_s Q_s E_g(s) \tag{30}$$

where the weights are

$$Q_s = \sum_F \overline{\Pi}_F(s)^{-1} \langle \overline{\Pi}_F(\sigma)\rangle . \tag{31}$$

Hence the band gap of the random alloy can be written as a sum over those of ordered structures with weights given by the thermal average in Eq. (31) (this was illustrated in Table VII). While in general the weights Q_s of Eq. (31) cannot be interpreted as probabilities, in special cases[4(a)] they can. This is the case where the maximum figure F_{max} defines a translational repeat unit such that the ordered structures $\{s\}$ can be described as a superposition of cells n each being a particular arrangement of A and B atoms at the vertices of F_{max}. A simple example is when F_{max} is a nearest-neighbor tetrahedron ($k=4$) and $\{s\}$

are the $2^k = 16$ ordered structures $A_n B_{4-n}$ ($0 \leq n \leq 4$) spanned by this figure.[4] Then, the alloy band gap can be written as

$$\langle E_g \rangle_R = \sum_n Q_n(x, T) E_g(n) . \qquad (32)$$

At $T \to \infty$, it can be shown[4(a),4(b)] that $Q_n(x, \infty)$ becomes the random Bernoulli probability.

To see the effect of SRO on the alloy's band gap, we use the function $Q_n(x, T)$ computed in Ref. 4(b) for $Al_{1-x}Ga_x As$ and contrast the band gap $\langle E_g \rangle_R$ obtained from Eq. (32) with that obtained by substituting in this equation the Bernoulli distribution $Q_n^{(R)}(x, \infty)$ appropriate to the perfectly random alloy with no SRO. To exaggerate the effects, we use $Q_n(x, T)$ at a rather low temperature of 300 K (well below growth temperatures), where $\Delta Q_n(x, T)$ of Eq. (28) is large. We can use for $E_g(n)$ in Eq. (32) the band gaps of the 16 structures that are exactly spanned by F_{max} being a tetrahedron. These consist of five distinct configurations A (zinc blende), A_3B (luzonite), AB (CA), AB_3 (luzonite), and B (zinc blende) with ratios 1:4:6:4:1. To obtain more accurate results (see Table VII) we can use the renormalization trick of Ref. 4 whereby $E_g(n)$ [denoted $\Delta E(n, V)$ in Eq. (2.13) of Ref. 4(a)] is replaced by its renormalized from [denoted $\bar{\varepsilon}_n$ in Eq. (2.27) of Ref. 4(a)]. Equation (6.16) of Ref. 4(a) gives this quantity where all pair interactions are folded in. Figure 4 shows the difference between the alloy band gap computed with SRO [i.e., using $Q_n(x, T = 300$ K)] and the band gap of the perfectly random alloy [using $Q_n(x, \infty)$]. We see that (i) SRO reduces the band gap in the GaAs-rich region and raises it in the AlAs-rich region but that (ii) the magnitude of the effect is totally negligible, even though we have used ΔQ_n at low temperatures where SRO effects are largest. Both observations disagree with the model of Ref. 55. We conclude that

SRO effects created by finite temperature equilibrium have but a negligible effect on the direct band gap of AlAs/GaAs alloys.

C. Predicting direct band gaps of ordered structures

The central challenge of a cluster expansion approach of the sort attempted here is to design a lattice configuration with a given band gap, or one that is likely to have the largest possible (or smallest possible) direct band gap at a given composition x. For systems with a large number of sites M per unit cell, small gap structures will, most likely, correspond to superlattices with a sufficiently large number of GaAs layers (so quantum confinement, that increases the gap, is small) and any number of AlAs layers. This is so because in long-period type-I superlattices the band gap equals that of the constituent with the smaller of the two band gaps. However, prediction of structures with a *maximum* band gap (or structures with minimum gap but with a small number of atoms per unit cell) is nontrivial. To the authors' knowledge, there is no way that this question can be addressed in the context of **k** space (i.e., band-structure) methods except by trial-and-error calculations of the band structure of many different structures. However, the establishment of a reasonably converged cluster expansion in terms of *linear* contributions [Eq. (13)] from real-space atomic clusters enables a direct and simple way for addressing this problem. Simply stated, the linear form states that cluster (k, m) contributes an energy $D_{k,m} \bar{\Pi}_{k,m}(s) p_{k,m}$ to the band gap of structure s. Figure 3(a) shows, for example, that the third-neighbor pair L_2 and the third-neighbor triangle L_3 tend to significantly increase (decrease) the band gap in structures having positive (negative) values for $\bar{\Pi}_{L2}$ and $\bar{\Pi}_{L3}$. This establishes a simple design principle for band gaps of substitutional systems in terms of a search for structures with a desired frequency $\bar{\Pi}_{k,m}$ of the appropriate atomic figures.

We have constructed a file containing all possible atomic configurations on a binary fcc lattice with up to M sites (avoiding duplication[57]), and obtained their $\bar{\Pi}_{k,m}$ values. To find the structures with maximum band gap, we then search $P(\sigma)$ of Eq. (13) (using our calculated interaction energies p_F) for all 2^M configurations σ. In this preliminary study we searched only $2^{10} = 1024$ configurations with up to $M = 10$ sites.

We found that for $M \leq 10$ atoms, all maximum band-gap structures are [201] superlattices. Of these, some were included (unwittingly), in our data base of structures (Table II), i.e., the $(AlAs)_1/(GaAs)_2$ "$\gamma 1$" structure at $x = \frac{2}{3}$, the $(AlAs)_2/(GaAs)_2$ chalcopyrite (CH) structure at $x = \frac{1}{2}$, and the $(AlAs)_2/(GaAs)_1$ "$\gamma 2$" structure at $x = \frac{1}{3}$. Unsuspected maximum gap structures that were not included in our data base but were identified in the search are the $(AlAs)_1/(GaAs)_4/(AlAs)_1/(GaAs)_2$ [201] superlattices at $x = \frac{1}{4}$, for which the ten-term cluster expansion gave a predicted direct band gap of 1.10 eV, and the $(AlAs)_4/(GaAs)_1/(AlAs)_2/(GaAs)_1$ [201] superlattice at $x = \frac{3}{4}$, for which the same cluster expansion predicts a
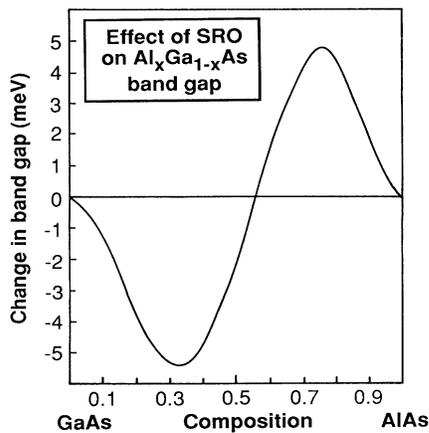

FIG. 4. Difference in the direct band gap of AlAs/GaAs alloys containing short-range order (calculated from the thermodynamic model of Ref. 4 at $T = 300$ K) and the perfectly random alloy ($T \to \infty$).

gap of 1.72 eV.

Although computationally expensive, we have calculated the self-consistent band structure of the former $x = \frac{1}{4}$ structure using the pseudopotential method with precisely equivalent basis sets and Brillouin-zone sampling used for all other structures. This gave a direct band gap of 1.07 eV, very close to the prediction of the cluster expansion of 1.10 eV. This highlights the usefulness of such cluster expansions in the design of materials with specified properties. Note that both values are larger than the linearly weighted average $\frac{1}{4}E_g$(AlAs) $+\frac{3}{4}E_g$(GaAs) and then the largest band gap at $x = \frac{1}{4}$ contained in our "basis set" (e.g., see Fig. 3).

VII. SUMMARY

We have demonstrated that the direct band gap of substitutional AlAs/GaAs systems can be usefully expanded in terms of contributions from a hierarchy of atomic clusters. We find that two-body and three-body figures are dominant, and that the expansion requires larger figures than needed to express the total energy surface. This expansion enables the prediction of the band gap of random alloys, alloys with short-range order, and gap of complex ordered structures.

[1]D. de Fontaine, in *Solid State Physics: Advances in Research and Applications,* edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1979), Vol. 34, p. 73; C. Domb, in *Phase Transitions and Critical Phenomena,* edited by C. Domb and H. S. Green (Academic, London, 1974), Vol. 3, pp. 357.

[2]J. Kanamori and Y. Kakehashi, J. Phys. (Paris) Colloq. **38,** C7-275 (1977); M. J. Richards and J. W. Cahn, Acta Metall. **19,** 1263 (1971); S. M. Allen and J. W. Cahn, *ibid.* **20,** 423 (1972).

[3](a) Z. W. Lu, S.-H. Wei, A. Zunger, S. Frota-Pessoa, and L. G. Ferreira, Phys. Rev. B **44,** 512 (1991); (b) Z. W. Lu, S.-H. Wei, and A. Zunger, Solid State Commun. **78,** 583 (1991); (c) Z. W. Lu, S.-H. Wei, and Z. Zunger, Phys. Rev. Lett. **66,** 1753 (1991); (d) S.-H. Wei, A. Mbaye, L. G. Ferreira, and A. Zunger, Phys. Rev. B **36,** 4163 (1987).

[4](a) L. G. Ferreira, S.-H. Wei, and A. Zunger, Phys. Rev. B **40,** 3197 (1989); (b) S.-H. Wei, L. G. Ferreira, and A. Zunger, *ibid.* **41,** 8240 (1990); (c) A. Mbaye, L. G. Ferreira, and A. Zunger, Phys. Rev. Lett. **58,** 49 (1987).

[5]J. W. D. Connolly and A. R. Williams, Phys. Rev. B **27,** 5169 (1983).

[6]K. Terakura, T. Oguchi, T. Mohri, and K. Watanabe, Phys. Rev. B **35,** 2169 (1987); S. Takizawa and K. Terakura, *ibid.* **39,** 5792 (1989).

[7]A. E. Carlsson, Phys. Rev. B **40,** 912 (1989).

[8]M. Sluiter, D. de Fontaine, X. Q. Guo, R. Podloucky, and A. J. Freeman, in *High-Temperature Ordered Intermetallic Alloys,* edited by C. T. Liu, A. I. Taub, N. S. Stoloff, and C. C. Koch, MRS Symposia Proceedings No. 133 (Materials Research Society, Pittsburgh, 1989), p. 3.

[9]J. M. Sanchez, F. Ducastelle, and D. Gratias, Physica A **128,** 334 (1984).

[10]D. B. Laks, R. Magri, and A. Zunger (unpublished) have shown how the randomization of the Ga and As sites on a diamond lattice changes the electronic structure and reduces the band gap of GaAs relative to the ordered zinc-blende phase.

[11]S. S. Batsanov, Izv. Akad. Nauk SSSR Neorg. Mater. **25,** 1292 (1989) [Inorg. Mater. **25,** 1091 (1990)].

[12]S. Froyen and M. L. Cohen, Phys. Rev. B **28,** 3258 (1983) showed in their calculations that GaAs metallizes in the sixfold coordinated NaCl structure.

[13]*Landolt-Börnstein Numerical Data and Functional Relationships* in Science and Technology, edited by O. Madelung, Landolt-Börnstein, Vol. 17f (Springer-Verlag, Berlin, 1987), pp. 103–122, collects data on Tl halides in the NaCl and in the CsCl structures, showing that minimum (indirect) gap in the CsCl structure is a few tenths of an eV smaller than the minimum (direct) gap in the NaCl structure.

[14]V. M. Glazov, S. N. Chizhevskaya, and N. N. Glayoleva, *Liquid Semiconductors* (Plenum, New York, 1969), pp. 101–131 have collected data showing that III-V semiconductors tend to metallize as their coordination number increases near the melting temperature.

[15]J. O. Dimmock, in *Proceedings of the 7th International Conference on II-VI Semiconducting* Compounds, Providence RI (Benjamin, New York, 1967), p. 277.

[16]In the wurtzite structure, the valence-band maximum is split into three components by the combined effect of spin-orbit and crystal-field effects. In increasing order of binding energies, these are $A = \Gamma_{9v}$, $B = \Gamma_{7v}$, and $C = \Gamma_{7v}$. The conduction band has Γ_{7c} symmetry. Light-polarized experiments hence reveal different transition energies for E∥c and E⊥c, where c is the hexagonal axis. In the zinc-blende structure, the valence-band maximum is split only by the spin-orbit effect into Γ_{8v} and Γ_{7v}, while the conduction band has Γ_{6c} symmetry. We consider in the text only transitions from the upper spin-orbit valence states.

[17]D. Theis, Phys. Status Solidi B **79,** 125 (1977).

[18]M. Cardona and G. Harbeke, Phys. Rev. **137,** A1467 (1965).

[19]Y. S. Park and F. L. Chan, J. Appl. Phys. **36,** 800 (1965).

[20]G. Hitler, thèse, Université de Paris VI, 1980; cited in Landolt-Börnstein *Numerical Data and Functional Relationships in Science and Technology* (Ref. 13), p. 126.

[21]M. Cardona, M. Weinstein, and G. A. Wolff, Phys. Rev. **140,** A633 (1965).

[22]E. Gutsche and J. Voight, in *Proceedings of the 7th International Conference on II-VI Semiconducting Compounds,* Providence, RI (Benjamin, New York, 1967), p. 337.

[23]V. V. Sobodev, V. I. Donetskina, and E. F. Zagainov, Fiz. Tekh. Poluprovodn. **12,** 1089 (1978) [Sov. Phys.—Semicond. **12,** 646 (1978)].

[24]We are unaware of experimental values for the band gap of CdSe in the zinc-blende structure.

[25]The indirect band gaps of SiC polytypes (that differ only in

third layer stacking) vary from 2.3 to 3.3 eV; see, e.g., N. L. Coleburn and J. W. Forbes, J. Chem. Phys. **48**, 555 (1968). However, these various gaps correspond to different folding **k** points in the parent Brillouin zone.

[26]D. Weaire and M. F. Thorpe, Phys. Rev. B **4**, 2508 (1971).

[27]J. D. Joannopoulos and M. L. Cohen, Phys. Rev. B **7**, 2644 (1973).

[28]F. M. Berkovskii, D. Z. Garbuzov, N. A. Goryunova, G. V. Loshakova, S. M. Ryvkin, and G. P. Shpenkov, Fiz. Tekh. Poluprovodn. **2**, 744 (1968) [Sov. Phys.—Semicond. **2**, 618 (1968)].

[29]M. A. Ryan, M. W. Peterson, D. L. Williamson, J. S. Frey, G. E. Maciel, and B. A. Parkinson, J. Mater. Res. **2**, 528 (1987).

[30]M. Cardona, T. Suemoto, N. E. Christensen, T. Isu, and K. Ploog, Phys. Rev. B **36**, 5906 (1987).

[31]M. Garriga, M. Cardona, N. E. Christensen, P. Lautenschlager, T. Isu, and K. Ploog, Phys. Rev. B **36**, 3254 (1987).

[32]S.-H. Wei and A. Zunger, Phys. Rev. B **39**, 3279 (1989).

[33]H. J. Lee, L. Y. Juravel, J. C. Woolley, and A. J. Spring-Thorpe, Phys. Rev. B **21**, 659 (1980); H. C. Casey, J. Appl. Phys. **49**, 3684 (1978).

[34]T. Fukui and H. Saito, Jpn. J. Appl. Phys. **23**, L521 (1984).

[35]A. Gomyo, T. Suzuki, and S. Iijima, Phys. Rev. Lett. **60**, 2645 (1988).

[36]B. T. McDermott, K. G. Reid, N. A. El-Masry, S. M. Bedair, W. M. Duncan, X. Yin, and F. H. Pollak, Appl. Phys. Lett. **56**, 1172 (1990).

[37]J. E. Bernard, S.-H. Wei, D. M. Wood, and A. Zunger, Appl. Phys. Lett. **52**, 311 (1988).

[38]S.-H. Wei and A. Zunger, Appl. Phys. Lett. **56**, 662 (1990).

[39]S.-H. Wei and A. Zunger, J. Appl. Phys. **63**, 5794 (1988).

[40]See Table II in R. Dandrea and A. Zunger, Phys. Rev. B **43**, 8962 (1991), where identical pseudopotential results and LDA corrections are used.

[41]C. Bosio, J. L. Staehli, M. Guzzi, C. Burri, and R. A. Logan, Phys. Rev. B **38**, 3263 (1988), and references therein.

[42]J. Ihm, A. Zunger, and M. L. Cohen, J. Phys. C **12**, 4409 (1979).

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

[44]G. Kerker, J. Phys. C **13**, L189 (1980).

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

[46]S. Froyen, Phys. Rev. B **39**, 3168 (1989).

[47]S. B. Zhang, M. S. Hybertsen, M. L. Cohen, S. G. Louie, and D. Tomanek, Phys. Rev. Lett. **63**, 1495 (1989); S. B. Zhang, M. L. Cohen, S. G. Louie, D. Tomanek, and M. S. Hybertsen, Phys. Rev. B **41**, 10058 (1990).

[48]A preliminary account of cluster expansion of the band gap of *random* alloys was given by A. Zunger, S.-H. Wei, L. G. Ferreira, and J. E. Bernard, Phys. Rev. Lett. **65**, 353 (1990); and S.-H. Wei, L. G. Ferreira, J. E. Bernard, and A. Zunger, Phys. Rev. B **42**, 9622 (1990).

[49]R. Magri, J. E. Bernard, and A. Zunger, Phys. Rev. B **43**, 1593 (1991).

[50]The D_4 structure has the lattice vectors $(0,1,1)$, $(1,0,1)$, and $(1,1,0)$. The atomic positions for the four Al atoms are $(0,0,0)$, $(\frac{1}{2},\frac{1}{2},0)$, $(\frac{1}{2},0,\frac{1}{2})$, and $(0,\frac{1}{2},\frac{1}{2})$. The positions of the four Ga atoms are $(1,0,0)$, $(1,\frac{1}{2},\frac{1}{2})$, $(\frac{1}{2},\frac{1}{2},1)$, and $(\frac{1}{2},0,-\frac{1}{2})$. The As atoms are in their regular tetrahedral sites.

[51]B. Koiller, M. A. M. Davidovich, and L. M. Falicov, Phys. Rev. B **41**, 3670 (1990); **42**, 3194(E) (1990), termed the D_4 structure as S_7 and the CP structure as S_7'. See discussion in R. Magri and A. Zunger, Phys. Rev. B **43**, 1584 (1991).

[52](a) B. Koiller, R. Osorio, and L. M. Falicov, Phys. Rev. B **43**, 4170 (1991); in their tight-binding calculations, these authors find $E_g(D4)-E_g(CP)=-0.023$ eV, close to our pseudopotential result of Eq. (26); (b) R. B. Capaz, B. Koiller, and J. P. von der Weid, Proc. Braz. School of Semiconductor Physics (São Paulo, 1991), in Revista Brasileira de Fisica (in press).

[53]A. B. Chen and A. Sher, Phys. Rev. B **23**, 5360 (1981).

[54]J. S. Nelson, A. F. Wright, and C. Y. Fong, Phys. Rev. B **43**, 4908 (1991).

[55]Y. Fu, K. A. Chao, and R. Osorio, Phys. Rev. B **40**, 6417 (1989).

[56]J. E. Hasbun, V. A. Singh, and L. M. Roth, Phys. Rev. B **35**, 2488 (1987).

[57]A. Zunger, S.-H. Wei, and L. G. Ferreira, Int. J. Supercomput. Applic. **5**, 34 (1991).