

Strain-Minimizing Tetrahedral Networks of Semiconductor Alloys

Jefferson Z. Liu, Giancarlo Trimarchi, and Alex Zunger

National Renewable Energy Laboratory, Golden, Colorado 80401, USA

(Received 13 May 2007; published 1 October 2007)

The atomic size mismatch between different binary semiconductors has been long known to limit their mutual solubility, leading instead to phase separation into incoherent phases, forming *inhomogeneous* mixtures that severely limit technological applications that rely on carrier transport. We show here that this atomic size mismatch can lead, under coherent conditions, to the formation of a *homogeneous* alloy with characteristic (201) two-monolayer ordering. This occurs because such specific layer arrangement corresponds to a unique strain-minimizing network in tetrahedral systems.

DOI: 10.1103/PhysRevLett.99.145501

PACS numbers: 61.66.Dk, 64.70.Kb

The atomic configuration in semiconductor alloys combining compounds, such as Si-Ge, InP-GaP, InN-GaN, or ZnO-ZnS, decides much of their electronic, optical, and transport properties [1,2]. For example, the optical band gap and transport effective masses of ordered alloys differ significantly from those of random alloys of the same chemical composition [1], and the carrier mobility in phase-separated alloys can be very different from that of ordered alloys [3]. It is now understood that such atomic arrangements often represent the minimum-strain atomic topologies formed under the constrained conditions present during growth [1,4–7]. For example, in epitaxial growth, the equilibrium *AC-BC* solid solubility is enhanced, because phase separation into substrate-coherent constituents (*AC* on substrate) + (*BC* on substrate) have a higher strain energy than phase ($A_{1-x}B_xC$ on substrate) [4,5] and are suppressed in favor of formation of the latter solid solutions. Another example is the surface-reconstruction-induced long-range ordering [1]: Surface reconstruction creates strain patterning in a few near-surface layers [6,7] and leads, in turn, to an energetic driving force for selective incorporation of the smaller (larger) of two alloy atoms at high (low) strain subsurface sites, leading to the formation of long-range ordered ($(AC)_1/(BC)_1$ CuPt-like phases [8].

In addition to the above atomic configurations that emerge in substrate-coherent, surface-exposed growth, there is an important question about the nature of the lowest energy structures of three-dimensional bulk semiconductor alloys. This question pertains, for example, to either substrate-disengaged (“relaxed”) films or to growth methods lacking exposed surfaces (e.g., traditional liquid or melt approaches). Long-standing tradition [2,5] suggests that the formation enthalpy of such freestanding isovalent semiconductor alloys

$$\Delta H(x, \sigma) = E(x, \sigma) - [(1-x)E_{AC}(a_{AC}) + xE_{BC}(a_{BC})] \quad (1)$$

is usually positive [1,9] for either random or ordered configurations (where E is the total energy and σ represents a specific assignment of atomic occupation to each

lattice site). Consequently, the semiconductor alloy thermodynamics is characterized by the “miscibility gap” in the composition (x)-temperature (T) phase diagram, below which the alloy separates into *AC*-rich and *BC*-rich phases at their own equilibrium lattice constants a_{AC} and a_{BC} . Such “incoherent” phase separation is often a showstopper for many technological applications because of enhanced scattering associated with statistical inhomogeneities. In this Letter, we show that, in contrast to such an incoherent effect, isovalent zinc blende alloys can exhibit ordered “coherent,” three-dimensional tetrahedral networks with much lower strain energies than both the equivalent random alloys and the coherently separated constituents (i.e., both constituents have the same lattice constant at the interface). Such homogeneous equilibrium structures are characterized by alternate atomic bilayers along the (201) crystallographic direction, akin to the $(AC)_2/(BC)_2$ (201) chalcopyrite (CH) structure. Our result can serve as a new mechanism to understand the spontaneous formation of such (201) superlattices [10–12] in a continuous alloy growth, especially when a free surface is not present during growth (e.g., liquid epitaxy [11]).

Identification of minimum-strain topologies in isovalent semiconductors.—We first represent the strain energy of zinc blende semiconductor alloys using a valence-force-field model (VFF) [13], in which the strain energy results from the deviation of the nearest-neighbor bond lengths from ideal values in the respective binary constituents and from the deviation of the bond angles from the ideal tetrahedral value 109.5° . It turns out [14] that a zero-strain configuration is topologically impossible in tetrahedral $A_{1-x}B_xC$ networks, so all such structures freeze-in some finite amount of strain. Exploring strain-minimizing structures in pseudobinary $(AC)_p(BC)_q$ tetrahedral networks of composition $x = \frac{q}{N}$ (with $N = p + q$) involves surveying a large configuration space, e.g., $\sim 1.05 \times 10^7$ structures for $N = 32$ at $x = 0.25$. In the past, only a tiny fraction of this configurational space was explored (e.g., Ref. [15] studied 5 configurations at $x = 0.25, 0.5,$ and 0.75). Here we will survey a significant portion of the configurational space, using exhaustive enumeration [16] as well as genetic-

algorithm [17] methods, for a set of III-V semiconductor alloys, finding rather surprising minimal strain configurations.

The global view of $\Delta H(x, \sigma)$ for zinc blende $\text{In}_{1-x}\text{Ga}_x\text{N}$ alloys spanning $\sim O(2^{20})$ configurations [Fig. 1(a)] is obtained by an exhaustive enumeration approach [16], in which each of the $\sim O(2^{20})$ lattice configurations is obtained by decorating the cation sublattice by either In or Ga atoms and then minimizing its energy by relaxing all atom positions without swapping atoms. To examine whether the restriction to $N = 20$ alters the identity of the strain-minimizing structures, the search was extended to $N = 32$, aided by the genetic algorithm [17]. We find that the minimal strain energy configurations are short-period superlattice structures $(\text{InN})_2/(\text{GaN})_2\text{-CH}$ and $(\text{InN})_3/(\text{GaN})_1\text{-famatinitite}$ [Figs. 1(b) and 1(c)] with layers oriented along the (201) crystal direction whether we search cells containing $N = 20$ or $N = 32$ cations. Furthermore, in Fig. 1(a), over 50% of the lowest energy structures are well-defined layered structures along the (201) direction [i.e., (201) superlattices].

To examine if the minimum-strain topologies found are characteristic only of the VFF functional, we applied first-principle local-density approximation (LDA) calculations for InN-GaN alloys. Since direct LDA calculations of $\sim O(2^N)$ configurations are impractical, a mixed-basis cluster expansion [18] total-energy Hamiltonian $\Delta H_{\text{CE}}(\sigma)$ is constructed instead. Within the cluster expansion method, one defines a configuration σ as a specific

atomic occupation on each lattice site by an A or a B atom (spin variable $s_i = -1$ or 1).

$$\Delta H_{\text{CE}}(\sigma) = J_0 + \frac{1}{N} \left[\sum_i J_i s_i + \sum_{i,j} J_{ij} s_i s_j + \sum_{i,j,k} J_{ijk} s_i s_j s_k + \dots \right] + \sum_{\mathbf{k}} \Delta E_{\text{CS}}(\hat{\mathbf{k}}, x) F(\mathbf{k}, \sigma), \quad (2)$$

where J_{ij}, J_{ijk}, \dots are two-body, three-body, etc., interaction energies and $F(\mathbf{k}, \sigma) = |S(\mathbf{k}, \sigma)|^2 e^{-|\mathbf{k}|^2} / 4x(1-x)$. The last term represents atomic size-mismatch effects, in which the *constituent strain energy* $\Delta E_{\text{CS}}(x, \hat{\mathbf{k}})$ is defined as the elastic energy cost to form lattice matching at the interface of two semi-infinite slabs AC and BC of orientation $\hat{\mathbf{k}}$ [18]. The interaction energies $\{J_{ij}, \dots\}$ are obtained by fitting of $\Delta H_{\text{CE}}(\sigma)$ to a set of first-principle calculated excess total energies $\{\Delta H_{\text{LDA}}(\sigma)\}$ [19]. A good fit is obtained with 56 input structures, 20 pairs, and 5 triplet and 1 quadruplet figures. The final statistical prediction error [20] is ≤ 1 meV/atom. We again search the lowest energy structures by the exhaustive enumeration method up to 20 cations per unit cell. This confirms that, at $x = 0.5$ and 0.25, CH and famatinitite are the lowest energy structures in InN-GaN alloys.

In light of the good accuracy of VFF in characterizing InN-GaN alloys, we apply it to five other III-V zinc blende alloys InP-GaP, InAs-GaAs, InSb-GaSb, InAs-InSb, and InAs-InP (Table I). Exhaustive enumerations are adopted with up to 20 cations per unit cell. CH, universally, has the minimum formation enthalpy at $x = 0.5$. The results are summarized in Table I and compared with the random alloy energies.

Why do (201) structures have the capacities to best accommodate strain in the tetrahedral network?—The formation energy per atom of the superlattice (SL) $(AC)_p/(BC)_p$ with layer orientation $\hat{\mathbf{k}}$ can be expressed as

$$\Delta H(p, \hat{\mathbf{k}}) = \frac{2I(p, \hat{\mathbf{k}})}{p} + \Delta E_{\text{CS}}(\hat{\mathbf{k}}), \quad (3)$$

where the constituent strain energy $\Delta E_{\text{CS}}(\hat{\mathbf{k}})$ of Eq. (2) is the long-period $p \rightarrow \infty$ limit of Eq. (3) and $I(p, \hat{\mathbf{k}})$ is the interfacial energy. Figure 2 shows these SL energies per

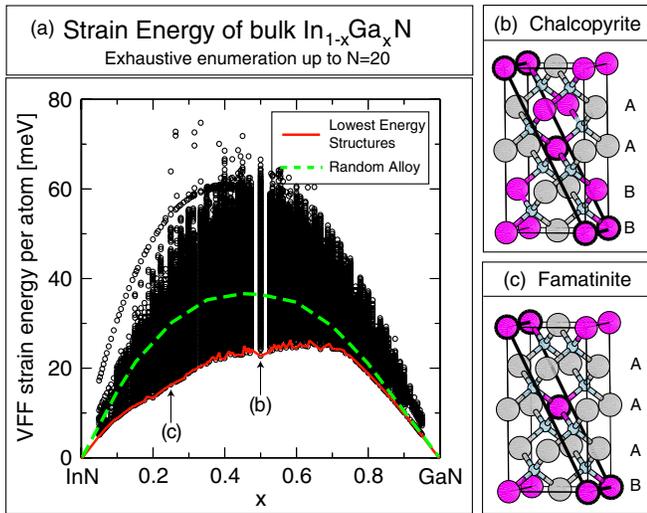


FIG. 1 (color online). (a) Formation enthalpy ΔH for $\text{In}_{1-x}\text{Ga}_x\text{N}$ calculated for all configurations with $N \leq 20$ [exhaustive enumeration (EE)] using valence-force-field relaxation. The red line connects all of the lowest energy structures at each composition, whereas the mixing energy of random alloy is drawn as the dashed green line. (b) CH and (c) famatinitite are predicted to be the lowest energy structures at $x = 0.50$ and $x = 0.25$ both by the EE method [16] ($N \leq 20$) and by the genetic-algorithm method [17] ($N = 32$), where the bold solid lines represent the [201] crystal planes.

TABLE I. Mixing enthalpy of the random alloy $\Delta H(R)$ at composition $x = 0.5$ and formation enthalpy of chalcopyrite ordered structure $\Delta H(\text{CH})$ for some III-V zinc blende alloys obtained with valence-force-field and LDA calculations [19] (in the parenthesis) in units of meV/atom. The lowest constituent strain $\Delta E_{\text{CS}}^{\text{min}}$ energies are also listed (see text).

| | InN GaN | InP GaP | InAs GaAs | InSb GaSb | InAs InSb | InAs InP |
|-------------------------------------|------------|------------|--------------|--------------|--------------|-------------|
| $\Delta H(R)$ | 37.3(40.6) | 20.0 | 15.9 | 11.3 | 12.4 | 3.00 |
| $\Delta E_{\text{CS}}^{\text{min}}$ | 38.7(39.1) | 21.9 | 17.5 | 12.4 | 14.1 | 3.45 |
| $\Delta H(\text{CH})$ | 22.8 | 13.2 | 10.5 | 7.52 | 8.20 | 1.92 |
| $\Delta H(\text{CH})$ | (19.6) | (8.63) | (6.60) | (4.47) | (8.58) | (2.94) |

atom obtained from VFF relaxation for $(\text{InN})_p/(\text{GaN})_p$ in different orientation $\hat{\mathbf{k}}$. The asymptotic $p \rightarrow \infty$ limit of each graph gives the numerical value of $\Delta E_{\text{CS}}(\hat{\mathbf{k}})$. In general, $\Delta E_{\text{CS}}(\hat{\mathbf{k}})$ is decided by the elastic constant in direction $\hat{\mathbf{k}}$ at $p \rightarrow \infty$, whereas $I(p, \hat{\mathbf{k}})$ depends on structural degrees of freedom for interfacial relaxation at short p . The emergence of the CH $p = 2$ (201) structure as the strain-minimizing network represents a favorable compromise between $\Delta E_{\text{CS}}(\hat{\mathbf{k}})$ and $I(p, \hat{\mathbf{k}})$.

(a) *Interfacial strain energy and symmetries (low p).*—Symmetry decides how many degrees of freedom are available for relaxation. The (111) and (100) interfaces allow, by symmetry, only a single degree of freedom for the common C atom to relax. Consequently, these superlattices exhibit small interfacial relaxation effects $I(p) \ll \Delta E_{\text{CS}}$, and the SL energy $\Delta H(p, \hat{\mathbf{k}})$ [Eq. (3)] is essentially p -independent [Fig. 2(a)]. In contrast, we find that the (101), (201), and (113) SLs have 2, 3, and 3 independent degrees of freedom, respectively, for the C atom to relax, and $I(2, \hat{\mathbf{k}})$ calculated by fitting the data of Fig. 2 to Eq. (3) are -24.2 , -24.3 , and -16.2 meV/atom, respectively. Consequently, $\Delta E(p, \hat{\mathbf{k}})$ for these directions shows a drop as the period p is reduced from $p = \infty$ (Fig. 2).

(b) *Coherent strain of thick SLs (large p).*—The $\Delta E_{\text{CS}}(\hat{\mathbf{k}})$ depend on elastic constants, and their orientation dependence relies on the epitaxial-strain-reduction factor [9]. It turns out that the (100) direction has the highest epitaxial-strain reduction and thus the lowest ΔE_{CS} . Consequently, the SLs with orientation at or near (100) directions, such as (104), (115), and (103), have small ΔE_{CS} (Fig. 2). But these low- ΔE_{CS} SLs at small p become identical to some other SLs which have either large $I(p, \hat{\mathbf{k}})$

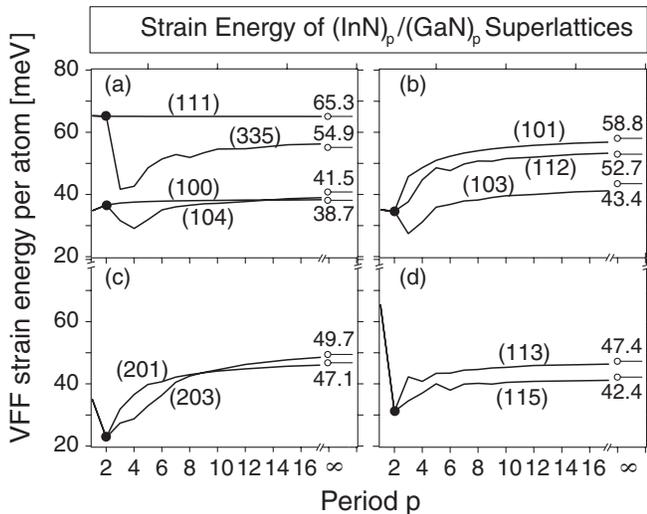


FIG. 2. Formation enthalpy of superlattices $(\text{InN})_p/(\text{GaN})_p$ calculated with the valence force field vs period p and layer orientation $\hat{\mathbf{k}}$. The constituent strain energies $\Delta E_{\text{CS}}(\hat{\mathbf{k}})$ ($p \rightarrow \infty$) are shown at the right. Note that (201) and (203) are identical structures at $p = 2$ and that (201) and (100) are identical structures at $p = 1$.

or large ΔE_{CS} . For example, the (104), (115), and (103) structures become (100), (113), and (101) SLs at $p = 2$ [21], respectively. Thus, CH is the structure with the lowest strain energy at $x = 0.5$.

Coherent ordering and phase separation.—Phase separation can be classified into two categories: incoherent [Fig. 3(a)] and coherent [Figs. 3(b) and 3(c)] phase separation [22]. In the incoherent case, AC and BC relax to their own equilibrium lattice constants a_{AC} and a_{BC} . This is the traditional phase separation realized by the high density of defects at the interface, such as dislocations, grain boundaries, etc. In the coherent case, after phase separation, AC and BC have the same lattice constant a_{\perp} at the interface. The energy cost to maintain coherence at interface $\hat{\mathbf{k}}$ is the constituent strain energy $\Delta E_{\text{CS}}(\hat{\mathbf{k}}, x)$ [Eq. (2)]. The lowest value $\Delta E_{\text{CS}}^{\text{min}}$ of $\Delta E_{\text{CS}}(\hat{\mathbf{k}}, x)$ with respect to $\hat{\mathbf{k}}$ is the energy of coherently separated phases. This can either raise the energy of the coherent constituents $AC(a_{\perp}) + BC(a_{\perp})$ above that of the lowest energy structures [Fig. 3(b)] or reduce the energy differences [Fig. 3(c)]. In Table I, $\Delta H(\text{CH})$ is always lower than the $\Delta E_{\text{CS}}^{\text{min}}$. This suggests that, if lattice coherence is maintained during bulk growth, CH will appear instead of phase separation at low temperature [Fig. 3(b)].

The ordering process has been simulated by the canonical Monte Carlo (MC) method for the InN-GaN alloy, using $\Delta H_{\text{CE}}(\sigma)$ as the energy functional. Figure 4(a) depicts the dependence of heat capacity and average energy on temperature at concentration $x = 0.5$. A coherent phase transition, where the disordered solid solution transforms to ordered CH structure, occurs at around 600 K. For comparison, we also calculate, from the same ΔH_{CE} , the incoherent phase-separation [Fig. 3(a)] temperature (using the “common-tangent-line” construction) 1870 K, which is about 3 times the coherent ordering temperature.

Short-range order fingerprints.—Since the ordering transition temperature ~ 600 K is below ordinary semiconductor growth temperature T_g , this first-order transition might not be detected directly in experiments. To facilitate

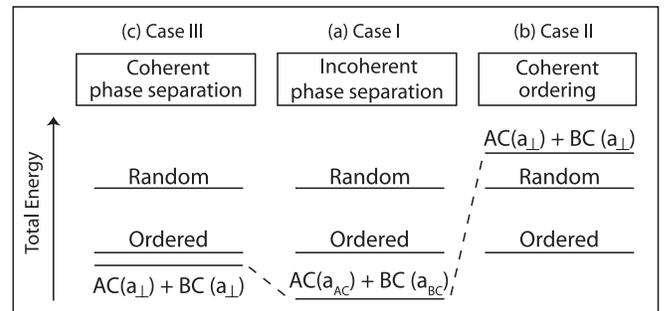


FIG. 3. The energetic order of various phases of $AC + BC$ leading to (a) incoherent (traditional) phase separation, where the separated phases AC and BC relax to their individual equilibrium lattice constants, and (b),(c) coherent structures, where AC and BC have the same lattice constant a_{\perp} at interface after phase separation.

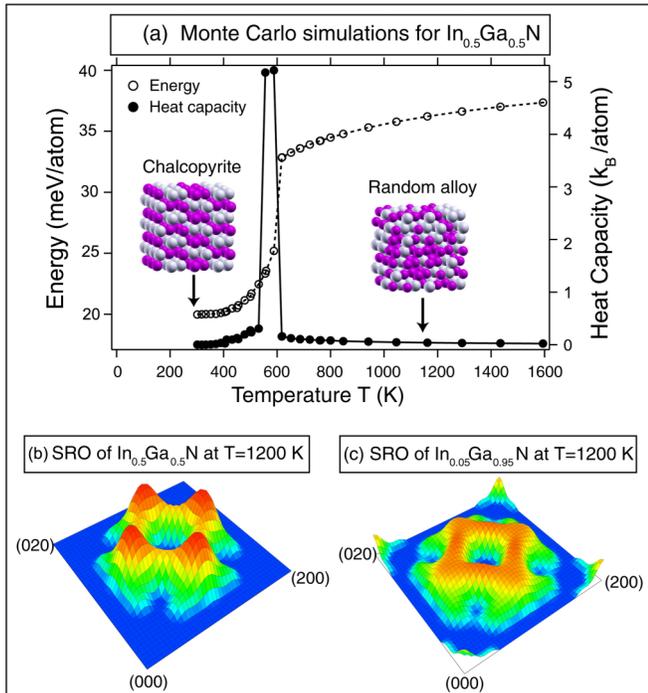


FIG. 4 (color online). (a) Heat capacity and average energy in MC simulations for an $\text{In}_{0.5}\text{Ga}_{0.5}\text{N}$ alloy. The two insets are atomic structures from MC simulation. Below T_c , the atomic structure is an $(\text{InN})_2/(\text{GaN})_2$ (201) superlattice, chalcopyrite. (b),(c) Fourier transformation of Warren-Cowley short-range order calculated from MC simulations.

the detection of such (201) ordered phases, we have calculated the Warren-Cowley short-range order (SRO) parameter, i.e., the deviation of pair correlations from the ideal random alloy [22]. Such deviation might be detected in the disordered solid solution *above transition temperature*. Ordering tendency is manifested by peaks in the SRO parameter away from $\mathbf{k} = \mathbf{0}$ in the reciprocal space, whereas peaks at $\mathbf{k} = \mathbf{0}$ indicate clustering or phase separation. Figures 4(b) and 4(c) show the calculated SRO for $\text{In}_{1-x}\text{Ga}_x\text{N}$ in reciprocal space at $T = 1200$ K and concentration $x = 0.50$ and 0.95 . The $x = 0.5$ alloy shows clear peaks at the $(\frac{1}{2}0)$ wave vectors, which correspond to the (201) superlattices CH at low temperature [Fig. 4(a)]. In contrast, for the alloy at $x = 0.95$, coherent *phase separation* has lower energy than the ordered structures [Fig. 3(c)]. The SRO peaks at $\mathbf{k} = \mathbf{0}$ show up. With decreasing temperature in our MC simulation, a phase-separation process is observed.

(201) CH ordering has been seen in bulk samples of InAs-GaAs [10,11] and GaAs-GaSb [12] alloys. The measured size of CH domains was quite small (≤ 50 nm), so it is reasonable to expect well-maintained coherence in such small size precipitates. Coherence in bulk samples, therefore, provides a new mechanism to understand the formation of CH.

In summary, a direct search of either the valence-force-field or the LDA-based functional identifies the (201)-

layered chalcopyrite motif as best able to accommodate atomic size-mismatch-induced strain in a tetrahedral network. Thus, under *coherent* growth conditions, chalcopyrite ordering provides lower energy structure than the phase separation. The bulk stable chalcopyritelike long-range ordering in semiconductor alloys is important [23], because it represents a reaction channel that could circumvent the detrimental incoherent phase separation.

G. T. acknowledges support by LDRD from NREL for work on searching lowest energy structures, and J.Z.L. and A.Z. were supported by the Office of Science of the DOE under Contract No. DOE-SC-BES-DMS on work for phase stability.

- [1] A. Zunger *et al.*, in *Handbook on Semiconductors*, edited by T. S. Moss (North-Holland, Amsterdam, 1992), Vol. 3b, p. 1399.
- [2] A.-B. Chen *et al.*, *Semiconductor Alloys* (Plenum, New York, 1995).
- [3] A. Chin *et al.*, *Appl. Phys. Lett.* **58**, 968 (1991).
- [4] A. Zunger *et al.*, *J. Cryst. Growth* **98**, 1 (1989); D.M. Wood *et al.*, *Phys. Rev. B* **40**, 4062 (1989).
- [5] G.B. Stringfellow, *J. Cryst. Growth* **27**, 21 (1974).
- [6] P.C. Kelires *et al.*, *Phys. Rev. Lett.* **63**, 1164 (1989).
- [7] J.E. Bernard *et al.*, *Phys. Rev. B* **44**, 11 178 (1991).
- [8] A. Gomyo *et al.*, *Phys. Rev. Lett.* **72**, 673 (1994).
- [9] R.G. Dandrea *et al.*, *Phys. Rev. Lett.* **64**, 36 (1990).
- [10] T.S. Kuan *et al.*, *Appl. Phys. Lett.* **51**, 51 (1987).
- [11] N. Nakayama *et al.*, in *GaAs and Related Compounds 1985*, edited by M. Fujimoto, *Inst. Phys. Conf. Ser. Vol. 79* (Hilger, London, 1986), p. 289.
- [12] H.R. Jen *et al.*, *Appl. Phys. Lett.* **48**, 1603 (1986).
- [13] P.N. Keating, *Phys. Rev.* **145**, 637 (1966).
- [14] J.L. Martins *et al.*, *Phys. Rev. Lett.* **56**, 1400 (1986).
- [15] J.E. Bernard *et al.*, *Appl. Phys. Lett.* **56**, 731 (1990).
- [16] L.G. Ferreira *et al.*, *Int. J. Supercomput. Appl.* **5**, 34 (1991).
- [17] G. Trimarchi *et al.*, *Phys. Rev. B* **74**, 014204 (2006).
- [18] D.B. Laks *et al.*, *Phys. Rev. B* **46**, 12 587 (1992).
- [19] The total energy is computed using the local density approximation [J.P. Perdew *et al.*, *Phys. Rev. B* **23**, 5048 (1981)] and the projector-augmented wave method, as implemented in the VASP [G. Kresse *et al.*, *Comput. Mater. Sci.* **6**, 15 (1996)] code. Choices of cutoff energy and k point mesh guarantee the energy convergence within 1 meV/atom.
- [20] The cross-validation score is calculated as an average prediction error for a given number of ordered structures by the cluster expansion that is least-square fitted from the rest of the configurations in a data pool.
- [21] Indeed, at $p = 1$, only (100) and (111) $(AC)_p/(BC)_p$ SLs are topological unique, and at $p = 2$, only five SLs (Fig. 2) are topological unique: (100), (111), (101), (201), and (311).
- [22] C. Wolverton *et al.*, *J. Phys. Condens. Matter* **12**, 2749 (2000).
- [23] $(AC)_2/(BC)_2$ chalcopyrite has a direct band gap if both AC and BC have a direct band gap [S.-H. Wei *et al.*, *Appl. Phys. Lett.* **56**, 662 (1990)].