Origin of the failed ensemble average rule for the band gaps of disordered nonisovalent semiconductor alloys

Jie Ma, Hui-Xiong Deng, Jun-Wei Luo, and Su-Huai Wei

Origin of the failed ensemble average rule for the band gaps of the disordered nonisovalent semiconductor alloys

Jie Ma,[1] Hui-Xiong Deng,[2, *] Jun-Wei Luo,[2] and Su-Huai Wei[1, †]

[1] *National Renewable Energy Laboratory, Golden, Colorado 80401, USA*
[2] *State Key Laboratory for Superlattices and Microstructures, Institute of Semiconductors, Chinese Academy of Sciences, P. O. Box 912, Beijing 100083, China*

Recent calculations show that the band gaps of the nonisovalent random alloys such as $Zn_{0.5}Sn_{0.5}P$ are much smaller than those of their ordered phases; that is, the band gap of the random alloy is not the ensemble averaged value of the ordered structures, in contrast to the trend observed in most isovalent semiconductor alloys and predicted by the cluster expansion theory. We show that this abnormal behavior is caused by the strong wavefunction localization of the band-edge states in the nonisovalent alloys. Moreover, we show that although the disordered phase of the isovalent alloys is similar to the random phase, for the nonisovalent alloy, the disordered phase deviates significantly from the random phase and the fully random phase is not achievable under the equilibrium growth conditions.

To broaden the range of the material properties for specific device applications, it is quite common to mix different elemental or binary semiconductor compounds to form alloys, because by varying the alloy composition and/or atomic configuration, the structural, electronic, transport, and optical properties of the alloys can be tuned[1–6]. The A_xB_yC type alloys can be classified into two categories: isovalent alloys, where A and B have the same valence state, such as $Ga_xIn_{1-x}P$, CdS_xTe_{1-x}, etc., and nonisovalent alloys, where A and B have different valence states, such as Cu_xIn_ySe, $Zn_{0.5}Sn_{0.5}P$, etc.. For the isovalent alloys, the compositions x and $y = 1-x$ usually vary smoothly and continuously in a wide range. However, for the nonisovalent alloys, the compositions can only exist around some discrete values, to satisfy the charge neutrality rule. For example, the Zn_xSn_yP, which has recently been proposed as a promising candidate for solar cells[7–9], exists only in a small range around $x = y = 0.5$; the Cu_xIn_ySe, which is widely used for thin-film solar cells[10,11], can also exist at other compositions besides $x = y = 0.5$, such as $CuIn_3Se_5$. The properties of the isovalent alloys have been extensively studied. For example, the band gap of the $Ga_xIn_{1-x}P$, which is an ideal material for the solid-state light-emitting diodes and high-efficiency multi-junction solar cells[12,13], can be tuned by varying the composition and ordering parameters. However, the properties of the nonisovalent alloys are poorly understood.

Isovalent and nonisovalent alloys can have some similar properties. For example, they all adopt ordered structures at low temperatures, and experience an order-disorder transition due to the increased entropy contribution as the temperature increases. They also have many different properties, even different chemical trends. For example, the ground states of free-standing isovalent alloys are generally phase separated, due to the positive mixing enthalpy, but those of the nonisovalent alloys are the ordered alloy structures satisfying the octahedral rule[14,15]. Their electronic structures can also be very

different. For the direct-band-gap isovalent alloys, it was found $E_g(CH) > E_g(CA) > E_g(CP)$[16], where CH, CA, and CP represent the ordered chalcopyrite, CuAu-, and CuPt-like structures respectively, and the random alloy has a band gap close to the ensemble averaged value of these ordered ones. However, for the nonisovalent alloys, the band gap of the random structure could be very small, and even smaller than that of the CP structure[9]. Nevertheless, such a small band gap has never been observed experimentally. The reason for this abnormal behavior is unclear. Because more and more nonisovalent alloys are currently considered for the optoelectronic applications, it is imperative to understand the general differences between the isovalent and nonisovalent alloys in their structural-functional relationships.
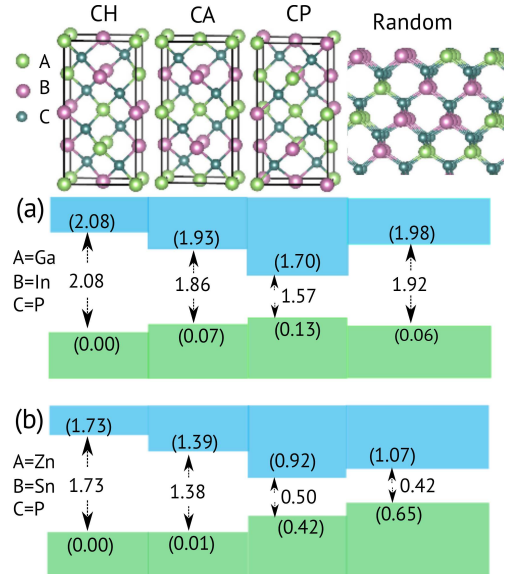

FIG. 1. The HSE band-edge energies of (a) the isovalent $Ga_{0.5}In_{0.5}P$ and (b) the nonisovalent $Zn_{0.5}Sn_{0.5}P$ alloys in the CH, CA, CP, and random (SQS) structures.

In this paper, using first-principles calculations, we compare the structural and electronic properties between the isovalent $Ga_{0.5}In_{0.5}P$ (for comparison we only consider $x = y = 0.5$) and the nonisovalent $Zn_{0.5}Sn_{0.5}P$ alloys. We have calculated the band-edge energies of the ordered CH, CA, CP, and fully random alloy structures. We find that for the isovalent alloys $E_g(CH) > E_g(CA) \sim E_g(Random) > E_g(CP)$, but for the nonisovalent alloys $E_g(CH) > E_g(CA) > E_g(CP) > E_g(Random)$. The abnormal trend of the nonisovalent alloys can be explained by the wavefunction localization of the band-edge states induced by the local charge transfer. We have also calculated the alloy phase-diagrams, and find that the disordered phase of the isovalent alloy is close to the random structure but a strong short-range ordering exists in the disordered nonisovalent alloy. The disordered nonisovalent alloy near the phase transition has been simulated, and the band gap agrees well with the experimental data. We also point out that for the nonisovalent alloys, a fully random phase is unachievable under the equilibrium growth condition.

Our calculations are based on the density functional theory[17,18] as implemented in the VASP code[19–21]. The projector augmented wave[22] pseudopotentials are employed, and the wavefunctions are expanded in a plane-wave basis with an energy cutoff of 400 eV. The random alloys are mimicked by the special quasirandom structures (SQS) in a 64-atom cell (32 mixed-atoms)[16,23]. We employ a $6 \times 6 \times 5$ k-point mesh for the CH structure, $6 \times 6 \times 4$ for the CA structure, $9 \times 9 \times 9$ for the CP structure, and $3 \times 7 \times 2$ for the SQS structure, respectively. The k-point meshes are fine enough to guarantee the convergence. For the electronic structure calculations, we employ the Heyd-Scuseria-Ernzerhof hybrid functional (HSE)[24–26]. We have also tested the generalized gradient functional (PBE)[27], and the calculated trends are similar although the band gaps are underestimated.
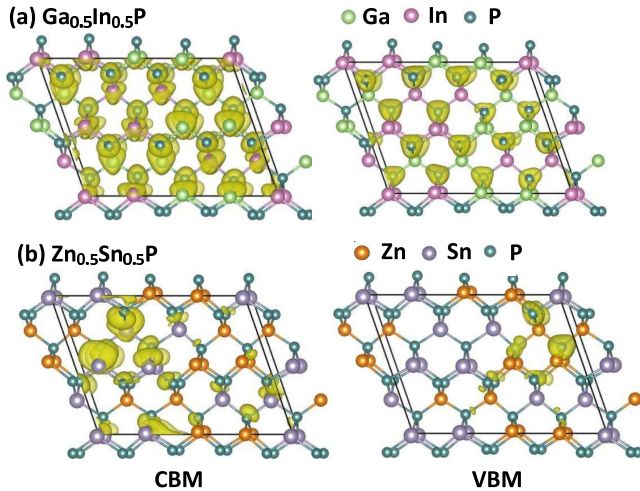

(a) $Ga_{0.5}In_{0.5}P$ Ga   In   P

(b) $Zn_{0.5}Sn_{0.5}P$ Zn   Sn   P

CBM   VBM

FIG. 2. The charge densities of the VBM and CBM in the random (a) $Ga_{0.5}In_{0.5}P$ and (b) $Zn_{0.5}Sn_{0.5}P$ alloys.

We start by comparing the band gaps of the ordered (CH, CA, and CP) and random (SQS) structures between the isovalent $Ga_{0.5}In_{0.5}P$ alloy and the non-isovalent $Zn_{0.5}Sn_{0.5}P$ alloy. The HSE band gaps and the relative positions of the conduction band minimum (CBM) and the valence band maximum (VBM) are shown in Fig. 1. The band edges are aligned using the core level of the P atoms. For the $Ga_{0.5}In_{0.5}P$, the CH structure has the largest band gap of 2.08 eV, the CP structure has the smallest band gap of 1.57 eV, and the CA structure has an intermediate band gap of 1.86 eV. The band gap of the random structure (1.92 eV) is close to the ensemble averaged value of the ordered structures, as one would expect. This trend is common in the isovalent alloys and has been successfully explained by the k-point foldings and level repulsions[16,28]. For the $Zn_{0.5}Sn_{0.5}P$, the band gaps of the CH, CA, and CP structures are 1.73, 1.38, and 0.50 eV, respectively. The trend of the band gaps of the ordered structures is similar to that of the $Ga_{0.5}In_{0.5}P$. However, surprisingly, the band gap of the random structure of the $Zn_{0.5}Sn_{0.5}P$ shows an abnormal trend: it is very small, even smaller than that of the CP structure.

To identify the origin of the different trends, it is helpful to compare the VBM and CBM energies of these ordered and random structures. As shown in Fig. 1, for both isovalent $Ga_{0.5}In_{0.5}P$ and nonisovalent $Zn_{0.5}Sn_{0.5}P$ alloys, from CH, CA, to CP, the VBM energy increases and the CBM energy decreases, so the band gap decreases. The band gap reduction from the CA to CP structure, however, is much larger for the $Zn_{0.5}Sn_{0.5}P$. For the $Ga_{0.5}In_{0.5}P$, the VBM energy changes less significantly than the CBM energy. This is because the $Ga_{0.5}In_{0.5}P$ is a common-anion alloy; therefore, the CBM, which is mostly derived from the cation s states, is more strongly affected by the alloying than the VBM, which is mostly derived from the anion p states. In the random structure, both the VBM and CBM energies are close to the averages of the VBM and CBM energies of the ordered structures, which confirms that the properties of the random alloy are the averages of those of the ordered structures. For the $Zn_{0.5}Sn_{0.5}P$, the CBM energy of the random structure, although low, is still close to the average of the CBM energies of the ordered structures. However, surprisingly, the VBM energy of the random structure is the highest among all the calculated structures, which is unexpected. Therefore, the band gap abnormality of the $Zn_{0.5}Sn_{0.5}P$ is mainly due to the unusual variation of the VBM energies.

The VBM and CBM charge densities of the random structure for the $Ga_{0.5}In_{0.5}P$ and $Zn_{0.5}Sn_{0.5}P$ are shown in Fig. 2. For the $Ga_{0.5}In_{0.5}P$, the VBM and CBM states are rather delocalized. However, for the $Zn_{0.5}Sn_{0.5}P$, the CBM localization is stronger, because the chemical potential difference between Sn and Zn is larger than that between In and Ga. However, because the CBM is mainly s state, which is delocalized, the CBM localization is still not too strong. In contrast, the VBM state of the random $Zn_{0.5}Sn_{0.5}P$ is strongly localized on the P atoms

surrounded by four Zn atoms. This is explained by the nonisovalent effect. Zn has two valence electrons and Sn has four. To form bonds, Sn transfers an electron to Zn, so that Zn is negatively charged and Sn is positively charged in the $Zn_{0.5}Sn_{0.5}P$. The strong Coulomb repulsion will push up the energies of the P p levels when the P atoms are surrounded by the negatively charged Zn atoms, and pull down the energies of the P p levels when the P atoms are surrounded by the positively charged Sn atoms. Furthermore, Zn has higher d orbitals than Sn, so the stronger $p - d$ coupling will also make the p level higher in energy when the P is surrounded by more Zn atoms. Therefore, the P atoms surrounded by four Zn atoms has the highest p level, and the VBM state is strongly localized on those sites.
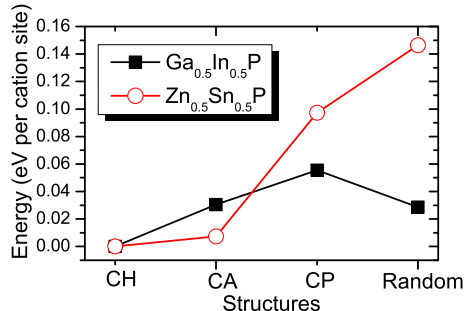

FIG. 3. The total energies per cation site of the ordered and random $Ga_{0.5}In_{0.5}P$ and $Zn_{0.5}Sn_{0.5}P$ alloys. The energies of the CH structure are set to zero.

For the $Zn_{0.5}Sn_{0.5}P$ in the CA and CH structures, all the P atoms are surrounded by two Zn and two Sn atoms, so the VBM energies of these two structures are similar. In the CP structure, half of the P atoms are surrounded by one Zn and three Sn atoms, and the other half are surrounded by one Sn and three Zn atoms. The VBM energy is mostly determined by the P atoms surrounded by three Zn atoms due to the Coulomb repulsion, so it is ~ 0.4 eV higher than those of the CA and CH structures, which explains the large band gap reduction from the CA to CP structure. In the random structure, the VBM energy is mostly determined by the P atoms surrounded by four Zn atoms due to the strongest Coulomb repulsion, so it is the highest among all the calculated structures, which explains the abnormal trend of the band gap.

Our analysis above clearly shows that for the isovalent $Ga_{0.5}In_{0.5}P$, the wavefunctions of the band-edge states are delocalized; therefore, their energies follow the ensemble averages of the ordered structures, as the cluster expansion theory predicts.[16] However, for the nonisovalent $Zn_{0.5}Sn_{0.5}P$, the wavefunctions of the band-edge states are localized, so their energies are determined by the local atomic configurations, not by the ensemble average of the lattice. This explains why the ensemble average rule fails for the nonisovalent alloys.

Not only the band gap but also the total energies of the nonisovalent alloys deviate from the ensemble average

rule. Figure 3 shows the calculated total energies of the ordered and random structures for the $Ga_{0.5}In_{0.5}P$ and $Zn_{0.5}Sn_{0.5}P$ alloys. For the $Ga_{0.5}In_{0.5}P$, the CH structure has the lowest total energy, the CP structure has the highest total energy, and the total energy of the CA is in between. The total energy of the random structure is again close to the ensemble averaged value of the ordered structures. This is the typical trend of the total energies for the isovalent alloys. However, for the $Zn_{0.5}Sn_{0.5}P$, the CH and CA structures have similar total energies, the energy of the CP structure is $\sim$0.1 eV per cation site higher, and the random structure has the highest energy (~ 0.14 eV per cation site higher than that of the CH structure). This abnormal trend is again due to the nonisovalent effect. In the nonisovalent alloys, to fulfill the octet rule, the local environment around a common anion atom should be charge neutral. For the CH and CA structures, because all the P atoms are surrounded by two Zn and two Sn atoms, the octet rule is satisfied and thus the total energies are similar and low. For the CP structure, because all the P atoms are either surrounded by one Zn and three Sn atoms or by one Sn and three Zn atoms, the octet rule is not satisfied, which raises the total energy. For the random structure, all types of the first-neighbor motifs exist. For the P atoms surrounded by Sn_4, Sn_3Zn, $SnZn_3$, and Zn_4, the octet rule is not satisfied, and thus the energy of the random structure is very high.

Now, we turn to the discussion of the thermodynamic properties of these two alloys. The alloy phase-diagrams are calculated using the cluster expansion approach as implemented in the ATAT code[29]. The cluster expansion coefficients are fitted to the energy calculated by the PBE functional. The equilibrium structures of the alloys at various temperatures are calculated by the Monte Carlo simulations in a 128000-atom cell.

We have calculated the phase diagrams for the coherently strained $Ga_{0.5}In_{0.5}P$ and $Zn_{0.5}Sn_{0.5}P$. The total energies per cation site of the equilibrium structures and the probabilities of the first-neighbor motifs of the P atoms: A_4, A_3B, A_2B_2, AB_3, and B_4, as a function of the temperature, are displayed in Fig. 4. In a fully random alloy, these probabilities are 0.0625, 0.25, 0.375, 0.25, and 0.0625, respectively, which are shown as the dashed lines in the figure. For the free-standing $Ga_xIn_{1-x}P$, the most stable phase at zero temperature is the phase-separated state: i.e., the alloy separates into GaP and InP. However, for the coherently (no broken bonds) strained $Ga_{0.5}In_{0.5}P$[30], where the lattice constant is fixed at the averaged value of GaP and InP, the CH structure is the most stable. At high temperatures, the entropy term plays a more important role and the disordered structures become more stable. An order-disorder phase transition occurs around $T = 350$ K for the coherently constrained $Ga_{0.5}In_{0.5}P$[31], where the total energies and the probabilities of the first-neighbor motifs of the P atoms change dramatically. After the phase transition, the total energies and the probabilities of the
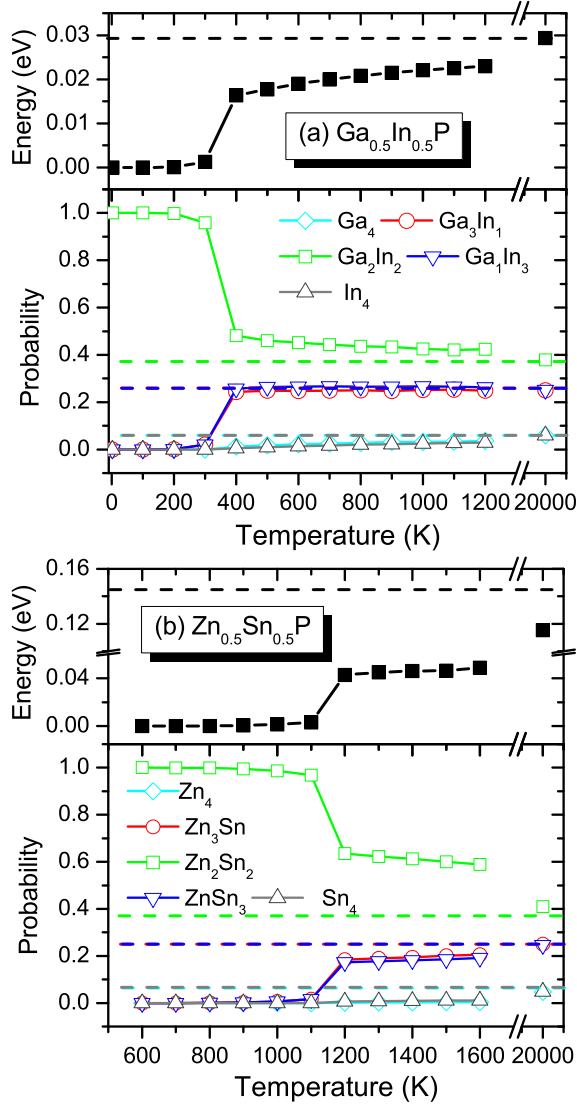
FIG. 4. The total energies per cation site and the probabilities of the first-neighbor motifs of the P atoms in the equilibrium structures as a function of the temperature for (a) the coherently strained $Ga_{0.5}In_{0.5}P$ on a lattice matched substrate, and (b) the $Zn_{0.5}Sn_{0.5}P$ alloys. The dashed lines indicate the corresponding total energies and probabilities in the random structure.

first-neighbor motifs of the P atoms are close to those of the random alloy, which indicates the disordered alloy is close to random. For the $Zn_{0.5}Sn_{0.5}P$, as the temperature increases, a phase transition occurs around $T = 1100$ K. However, different from the isovalent alloy case, the total energies are still ~ 0.1 eV per cation site lower than that of the random alloy after the phase transition. Similarly, the probabilities of the first-neighbor motifs of the P atoms are also quite different from those in the random alloy. Specifically, the probability of the P atoms surrounded by two Zn atoms and two Sn atoms is much higher, and the probabilities of the other four types are much lower, indicating that even after the phase tran-

sition, a strong short-range ordering still exists in the $Zn_{0.5}Sn_{0.5}P$ alloy. This is because for the nonisovalent alloys, the local motifs A_4, B_4, A_3B, and AB_3 have much higher energies, as discussed before, so the alloy tends to suppress the existence of those motifs even at the disordered phase. To study the disordered $Zn_{0.5}Sn_{0.5}P$ alloy, we have built a SQS that has the same correlation functions as the calculated disordered alloy at 1200 K. In this case, we find that the calculated probabilities of the first-neighbor motifs Zn_4, Zn_3Sn, Zn_2Sn_2, $ZnSn_3$, and Sn_4 are 0, 0.1875, 0.625, 0.1875, and 0, respectively, significantly different from the values in the random alloy. The calculated band gap of this disordered alloy is ~ 1.2 eV, which is in good agreement with the experiment[32]. As the temperature increases further, the total energies and the probabilities of the motifs change only slowly towards the values in the random alloy. We find that even at an unrealistic high temperature ($T = 20000$ K), the calculated total energy of the disordered alloy is still more than 20 meV per cation site lower than that of the random alloy, and the probabilities of the first-neighbor motifs of the P atoms have not achieved the values in the random alloy. These results suggest that for the nonisovalent alloys, the random alloy is not achievable under the equilibrium growth conditions in the experiments.

In summary, we show that the wavefunctions of the VBM and CBM states of the nonisovalent alloys are highly localized. This strong localization causes these states sample only a particular region in the lattice, not the whole lattice, which explains the failure of the ensemble average rule that is observed in the isovalent alloys. We find that for the isovalent alloys $E_g(CH) > E_g(CA) \sim E_g(Random) > E_g(CP)$, but for the nonisovalent alloys $E_g(CH) > E_g(CA) > E_g(CP) > E_g(Random)$. Moreover, we show that the disordered structures of the isovalent alloys are close to random, but for the nonisovalent alloys the random phase is not achievable under the equilibrium growth conditions. These findings are important in designing new optoelectronic devices based on the nonisovalent alloys.

[*] The first two authors contributed equally; hx-deng@semi.ac.cn

[†] swei@nrel.gov

[1] A. Hazarika, A. Layek, S. De, A. Nag, S. Debnath, P. Mahadevan, A. Chowdhury, and D. D. Sarma, Phys. Rev. Lett. **110**, 267401 (2013).

[2] S. S. Schmidt, D. Abou-Ras, S. Sadewasser, W. Yin, C. Feng, and Y. Yan, Phys. Rev. Lett. **109**, 095506 (2012).

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

[4] S. Wang and L.-W. Wang, Phys. Rev. Lett. **104**, 065501 (2010).

[5] S. Wei and A. Zunger, Phys. Rev. Lett. **61**, 1505 (1988).

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

[7] K. Nakatani, T. Minemura, K. Miyauchi, K. Fukabori, H. Nakanishi, M. Sugiyama, and S. Shirakata, Jpn. J. Appl. Phys. **47**, 5342 (2008).

[8] P. St-Jean, G. A. Seryogin, and S. Francoeur, Appl. Phys. Lett. **96**, 231913 (2010).

[9] D. O. Scanlon and A. Walsh, Appl. Phys. Lett. **100**, 251911 (2012).

[10] P. Jackson, D. Hariskos, E. Lotter, S. Paetel, R. Wuerz, R. Menner, W. Wischmann, and M. Powalla, Prog. Photovolt: Res. Appl. **19**, 894 (2011).

[11] W. N. Shafarman and L. Stolt, in *Handbook of Photovoltaic Science and Engineering*, edited by A. Luque and S. Hegedus (Wiley, Chichester, UK, 2003).

[12] R. R. King, D. C. Law, K. M. Edmondson, C. M. Fetzer, G. S. Kinsey, H. Yoon, R. A. Sherif, and N. H. Karam, Appl. Phys. Lett. **90**, 183516 (2007).

[13] J. F. Geisz, D. J. Friedman, J. S. Ward, A. Duda, W. J. Olavarria, T. E. Moriarty, J. T. Kiehl, M. J. Romero, A. G. Norman, and K. M. Jones, Appl. Phys. Lett. **93**, 123505 (2008).

[14] J. L. F. Da Silva, A. Walsh, and H. Lee, Phys. Rev. B **78**, 224111 (2008).

[15] J. L. F. D. Silva, J. Appl. Phys. **109**, 023502 (2011).

[16] S.-H. Wei, L. G. Ferreira, J. E. Bernard, and A. Zunger, Phys. Rev. B **42**, 9622 (1990).

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

[18] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[19] G. Kresse and J. Hafner, Phys. Rev. B **47**, 558 (1993).

[20] G. Kresse and J. Hafner, Phys. Rev. B **48**, 13115 (1993).

[21] G. Kresse and J. Furthmuller, Comput. Mater. Sci. **6**, 15 (1996).

[22] G. Kresse and D. Joubert, Phys. Rev. B **59**, 1758 (1999).

[23] A. Zunger, S.-H. Wei, L. Ferreira, and J. Bernard, Phys. Rev. Lett. **65**, 353 (1990).

[24] J. Heyd, G. E. Scuseria, and M. Ernzerhof, J. Chem. Phys. **118**, 8207 (2003).

[25] J. Paier, M. Marsman, K. Hummer, G. Kresse, I. C. Gerber, and J. G. Ángyán, J. Chem. Phys. **124**, 154709 (2006).

[26] We increase the Hartree-Fock portion to 28% for the $Ga_{0.5}In_{0.5}P$ calculations to fit the experimental band gaps.

[27] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[28] S.-H. Wei, L. G. Ferreira, and A. Zunger, Phys. Rev. B **45**, 2533 (1992).

[29] A. van de Walle, M. Asta, and G. Ceder, Calphad **26**, 539 (2002).

[30] S.-H. Wei, L. G. Ferreira, and A. Zunger, Phys. Rev. B. **41**, 8240 (1990).

[31] It should be noted that we calculate the phase diagramm of the coherently strained $Ga_{0.5}In_{0.5}P$ here. For the free-standing $Ga_{0.5}In_{0.5}P$ alloy, our calculated phase-transition temperature is around 800K.

[32] 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).