



# CHORUS

This is the accepted manuscript made available via CHORUS. The article has been published as:

## Global electronic structure of semiconductor alloys through direct large-scale computations for III-V alloys

$\text{Ga}_{\{x\}}\text{In}_{\{1-x\}}\text{P}$

Yong Zhang and Lin-Wang Wang

Phys. Rev. B **83**, 165208 — Published 26 April 2011

DOI: [10.1103/PhysRevB.83.165208](https://doi.org/10.1103/PhysRevB.83.165208)

# Insights into the global electronic structure of semiconductor alloys through direct large scale computations for III-V alloys $\text{Ga}_x\text{In}_{1-x}\text{P}$

Yong Zhang<sup>1\*</sup> and Lin-Wang Wang<sup>2</sup>

<sup>1</sup>*Department of Electrical and Computer Engineering and Center for Optoelectronics, University of North Carolina at Charlotte, Charlotte, NC 28223*

<sup>2</sup>*Lawrence Berkeley National Laboratory, Berkeley, CA 94720*

We critically examine two nominally equivalent approaches for treating a random alloy: (1) using one very large supercell as a direct simulation of the alloy, and (2) performing configuration averaging over many smaller supercells; and the common practice using a virtual-crystal as the reference for analyzing the alloy band structure and discussing the electronic transport in the alloy. Specifically, (1) we show that in practice the size of the “very large” supercell depends on the particular property of interest, and the ideal of the configuration averaging is only useful for certain properties; (2) we examine the assumed equivalency by comparing the results of the two approaches in bandgap energy, energy fluctuation, and inter-valley and intra-valley scatterings, and conclude that the two approaches often lead to non-equivalent physics; (3) by using a *generalized moment method* that is capable of computing the global electronic structure of a sufficiently large supercell (e.g.,  $\sim 260,000$  atoms), we are able to obtain the intrinsic broadening of a  $\Gamma$ -like electron state, caused by the “inelastic” intra-valley scattering, in a direct bandgap semiconductor alloy; (4) we demonstrate an efficient way to construct the effective dispersion curves of the alloy, with high accuracy for calculating effective masses and examining anisotropy and non-parabolicity of the dispersion curve; and (5) finally, we discuss the limitation of using virtual crystal approximation as the reference for evaluating the alloy scattering and studying the transport property.

## I. Introduction

Disordered alloys have been traditionally treated by approximate methods such as *virtual crystal approximation* (VCA) or *coherent potential approximation* (CPA).[1, 2] However, there could be serious omissions in such approximate methods. In VCA, for an alloy  $A_xB_{1-x}C$ , the potentials of A and B substitutional atoms have been averaged to yield a simple periodic system over the primitive unit cell. In CPA, the energy dependent transfer matrix (t-matrix) of atoms A and B has been averaged. Correspondingly, an energy dependent coherent potential is introduced in the Hamiltonian, which again makes the potential periodic over the primitive unit cell, albeit the resulting Hamiltonian is non Hermitian. This periodic, but non-Hermitian Hamiltonian gives us a band structure with a finite width for each band state. However, the multiple scatterings between A and B have been ignored in such a treatment. In order to include all the scattering effects in a straightforward manner, another approach is to calculate directly a finite disordered system. One such method is the special quasi-random structure (SQS) method,[3] where a relatively small superlattice system is constructed to have the atom-atom correlation functions as close in the true random system as possible. But the SQS method has a major limitation, that is, a SQS has a well defined symmetry that has undesirable physics consequence (for instance, the splitting of the valence band that is supposed to be degenerate for a random alloy). With the increased computer power, it has become possible to use a very large supercell to represent the true random alloy. However, in the past, the use of a very large supercell (in the order of one million atoms), with the help of a folded spectrum method, typically allowed calculating only a few states near a selected energy level, such as the conduction band minimum or valence band maximum.[4] As demonstrated in the current paper, we can now use a supercell containing quarter million atoms (or more) to study the alloy electronic structure for practically any required energy spectral range, with the help of a more efficient computational technique. This opens up a

different avenue to investigate the electronic structure of either totally random or partially ordered alloys. In random systems, fluctuation and statistical averaging become the essential theme. The fluctuation in a finite size system, which is only available in the supercell approach, can in principle provide information that can be compared directly with the result of optical spectroscopy. In this paper, we will discuss the different statistical approaches, and the meanings of the differently averaged results. We will also show how the supercell approach can yield the alloy band structure and band width, and how they are related to the CPA band structure and the roles of inter-valley and intra-band scatterings when using the VCA as a reference.

By definition, an alloy does not have translational symmetry. In principle, we need to deal with a quantum mechanical system with a large number of atoms in any macroscopic device involving the alloy. For instance, a cube of 100 nm size contains  $> 40$  million atoms. If one is not constrained by the computational power, what would be the appropriate structure size to simulate the electronic structure of an alloy? Intuitively, two approaches can be used to solve the quantum mechanical problem of the alloy:[1] (1) Constructing and solving one single sufficiently large structure as representative for the alloy; (2) Constructing many relatively small size structures (but still containing many atoms, much larger than any typical SQS), then solving them individually followed by configuration averaging. It is generally believed that these two approaches should in principle lead to the same results if the “small” size system is large enough, because a macroscopically large crystal may be visualized as composed of many smaller but different pieces.[1] However, such equivalency has never been rigorously examined. Even though the CPA theory is generally formulated based on configuration averaging,[1, 2] because the single site approximation is typically used in a real calculation, the result becomes independent of the size and detail of the configurations. The supercell approach makes it feasible

to directly test these two approaches. However, when actually implementing either approach, one will encounter a few practical issues. For instance, what structure size can be considered as “sufficiently large” to be representative for the alloy and what size is appropriate to be used as the “small structure” for the purpose of configuration averaging? If the size is too small, will the averaged result deviate from that of the infinitely large supercell result? Does the fluctuation of the finite size results have any physical meaning? Perhaps the two most basic pieces of knowledge regarding the electronic structure of a semiconductor alloy are its fundamental bandgap and inhomogeneous broadening. Therefore, in this work, we will examine the two approaches focusing primarily on these two basic physical properties. We find that greatly different supercell sizes are required to address different alloy problems. For instance, one supercell size that is capable of yielding a satisfactory accuracy in the bandgap might be totally inadequate or result in misleading information on the alloy scattering. Thus, it is neither practical nor necessary to find one supercell size that suits all purposes. We will therefore discuss what the appropriate supercell size is for a particular property of interest.

In the supercell simulation of the alloy, a practical issue is the selection of a boundary condition. There are three possible boundary conditions that can be applied in the calculation: (1) periodic repetition of the supercell, (2) an isolated supercell in vacuum, and (3) periodic repetition but with barriers inserted between the alloy structures. Unfortunately, none of them is a perfect choice. The option (1) will lead to the coupling among the supercell units, option (2) surface states, whereas option (3) the coupling between the alloy and the barrier material. Option (1) is the easiest way to implement, and one may hope that with the use of a large enough supercell the inter-supercell coupling becomes negligible. There is another potential advantage of choosing (1), that is, the periodic system allows the use of  $\mathbf{k}$  points (although in a very small

Brillouin zone for a large supercell). As we will see later, the dispersion of the supercell can be connected with the alloy band structure in a seamless fashion for a large supercell.

Firstly, the required supercell size depends on our convergence goal. A reasonable convergence goal for the electronic structure calculation of the semiconductor alloy is to achieve the accuracy obtainable by optical spectroscopy that is usually more accurate than other experimental techniques. For a semiconductor alloy that is well prepared and characterized (e.g., the composition is accurately known), the bandgap  $E_g$  can be determined with an accuracy of a few meV for a given composition  $x$  by optical spectroscopy, for instance, for the prototype system  $\text{Ga}_x\text{In}_{1-x}\text{P}$  to be examined in this work.[5] A benchmark testing has revealed that for  $\text{Ga}_x\text{In}_{1-x}\text{P}$  alloy, the bandgap converges to about 10 meV or about 1 meV by configuration averaging over 432-atom supercells or 3456-atom supercells, when comparing to the result of one 27648-atom supercell.[6] In this work, by applying a more efficient technique, *generalized moment method* (GMM), we are able to calculate an even larger supercell with 259200 atoms and further confirm the convergence of the configuration average of the 3456-atom supercell for the bandgap energy. Note that the use of the GMM does not affect the convergence of a particular material property with respect to the supercell but the efficiency to reach the desirable accuracy. Therefore, for the purpose of calculating the alloy bandgap, we have an established guideline for the supercell size for the two approaches: 3 – 4 thousand atoms for doing the configuration averaging or around 30 thousand atoms for using one representative configuration. Although this guideline is observed based on the work for  $\text{Ga}_x\text{In}_{1-x}\text{P}$ , it is likely valid for most conventional semiconductor alloys. For a linear scaling method like the GMM or the folded spectrum method, the fact that the supercell size differs by a factor less than 10 for the two approaches suggests that the configuration averaging approach does not necessarily require less

computational effort, considering the needs to calculate many configurations (typically 50-100 configurations) for the small system to provide good statistics. Secondly, the required supercell size depends on the specific alloy property to be studied. For instance, to investigate the intrinsic spectral broadening of a band edge alloy state, the approach of using one very large supercell is necessary, although the statistical fluctuation of many smaller structures might have its own physical significance.

One can compute directly a number of electronic and optical properties, such as bandgap, density of states, dielectric functions, for a semiconductor alloy without having to introduce a reference Hamiltonian. However, for the understanding of the spectral broadening due to the alloy fluctuation and closely related transport properties, it is useful to introduce a reference Hamiltonian, typically the VCA, to provide the reference states free of alloy scatterings.[1, 2] A spectral function  $A_n(\mathbf{k}, E)$  can be defined as:

$$A_n(\mathbf{k}, E) = \sum_i |\langle \psi_i | \phi_{n,\mathbf{k}} \rangle|^2 \delta(E - \varepsilon_i) \quad (1)$$

here  $\psi_i$  is the alloy eigen state with energy  $\varepsilon_i$ , and  $\phi_{n\mathbf{k}}$  is the VCA wavefunction at a  $\mathbf{k}$  point of the primary cell Brillouin zone (BZ) with a band index  $n$ . This spectral function describes how a VCA state is decomposed into different alloy states. In a way it is a  $\mathbf{k}$ -resolved density of states of the alloy. If the alloy fluctuation does not introduce any coupling among the VCA states, the spectral function would be a single  $\delta$  function in energy  $E$  for a given  $\mathbf{k}$ . For a disordered system, the finite width of  $A(\mathbf{k}, E)$  can be related to the quasi-particle life time. It is very similar to the finite energy width in the CPA theory, and also related to the scattering rate based on the VCA treatment. The spectral width for a general  $\mathbf{k}$  point of the VCA band structure, high symmetry points of the BZ in particular, will be an important topic to be discussed in this paper. Note that

in the direct supercell calculation, the spectral width can be induced by elastic coupling (among degenerate  $\mathbf{k}$  points) or inelastic coupling (among non-degenerate  $\mathbf{k}$  points). In most CPA and VCA treatments, only the elastic coupling is considered, although the second order perturbation theory involving non-degenerate states is used to explain the band gap bowing effect based on the VCA. As a result, the spectral width is simply proportional to  $\rho(E)$ , the density of states of the unperturbed reference system (e.g., the VCA system). We will show that, in the full supercell treatment, this is not generally true.

A number of semiconductor alloys have been previously investigated for their spectral functions  $A(\mathbf{k},E)$ . For III-V alloys with relatively small lattice and chemistry mismatch, such as  $\text{Al}_x\text{Ga}_{1-x}\text{As}$  and  $\text{Ga}_x\text{In}_{1-x}\text{As}$ , using the CPA or Molecular-CPA theory,[7, 8]  $A(\mathbf{k},E)$  is found to be very sharp, practically a  $\delta$ -function, in the direct bandgap composition region at  $\mathbf{k} = 0$  (the  $\Gamma$  point), because the elastic inter-valley scattering is not possible and the inelastic intra-valley scattering is very weak for the lowest conduction band state, which explains why the electronic transport in a direct bandgap alloy usually does not degrade drastically from a true crystalline structure. For the indirect bandgap alloy  $\text{Si}_x\text{Ge}_{1-x}$ , the spectral width of  $A(\Gamma,E)$  were found to vary from 1 meV at  $x = 0.1$  to 0.2 eV at  $x = 0.5$ . Because the alloy bandgaps are indirect, the  $\Gamma$ -like state is in resonance with other  $\mathbf{k}$  points that are degenerate with the  $\Gamma$  state, leading to significant inter-valley coupling.[9] We have recently found that, by a direct calculation using the 27648-atom supercell, for  $\text{Ga}_x\text{In}_{1-x}\text{P}$  with  $x = 0.8$  in the indirect bandgap region the  $A(\Gamma,E)$  spectrum shows barely a peak to allow for the identification of the  $\Gamma$ -like state in the alloy.[10] However, for  $x = 0.5$  in the direct bandgap region, the 27648-atom supercell is far from adequate to reveal the intrinsic  $A(\Gamma,E)$  spectrum, because the alloy scattering to the  $\Gamma$  state is very weak. Very recent extension of the similar analysis to more strongly mismatched alloys  $\text{In}_x\text{Ga}_{1-x}\text{N}$ ,

using the supercell method with size up to about 4,000 atom,[11] has yielded qualitatively the same observation as for the other simpler alloys. By calculating  $A(\mathbf{k},E)$  throughout the BZ, one can construct an effective band structure for an alloy using a properly defined VCA reference band structure, which has been done for numeral semiconductor alloys:  $\text{Hg}_{1-x}\text{Cd}_x\text{Te}$ , [12]  $\text{Si}_x\text{Ge}_{1-x}$ , [9],  $\text{Ga}_{0.5}\text{In}_{0.5}\text{As}$  and  $\text{ZnSe}_{0.5}\text{Te}_{0.5}$ , [7] and the latest  $\text{In}_x\text{Ga}_{1-x}\text{N}$ . [11] It turns out that the most challenging task is to accurately calculate  $A(\Gamma, E)$  for a direct bandgap material, which has never been explicitly calculated using a sufficiently large supercell, but is fundamentally important for the understanding of the alloying effect.

In this work, we first in “Section II - Theoretical Methods” introduce the *generalized moment method* that can evaluate the spectral function  $A(\mathbf{k}, E)$  far more efficiently than previously used approaches, particularly for very large supercells; then in “Section II – Results and Discussions”, we apply this method to the prototype alloy system  $\text{Ga}_x\text{In}_{1-x}\text{P}$  to investigate the intra-valley scattering of the  $\Gamma$  point or the inhomogeneous broadening of the single electron state (II-1), analyze the inter-valley and intra-valley scatterings at the L and X points and the alloy scattering induced spectral broadening for the general alloy states far away from the band edge, construct the effective dispersion curves and calculate the effective masses of the alloy bands (II-2), for  $x = 0.5$  in the direct bandgap region; compare the spectral functions obtained by configuration averaging with moderate size supercells and one very large supercell for  $x = 0.8$  in the indirect bandgap region (II-3); and discuss the limitations of studying alloy scattering within the framework of VCA. Finally, a summary is given in Section III.

## II. Theoretical methods

An empirical pseudopotential method (EPM), [13] which has been successfully applied to the study of the electronic structure of  $\text{Ga}_x\text{In}_{1-x}\text{P}$  alloys, [6, 10] is used in this work. The (reciprocal-

space) empirical pseudopotential takes the form of  $v(q, \boldsymbol{\epsilon}) = v(q, 0)[1 + a_s \text{Tr}(\boldsymbol{\epsilon})]$ , where  $v(q, 0)$  is the value at equilibrium lattice constant, and  $\text{Tr}(\boldsymbol{\epsilon})$  is the trace of the local strain (approximated by the relative change of the tetrahedron volume).[13] The extra term associated with strain parameter  $a_s$  offers a significant improvement over the conventional EPM in the presence strain and lattice relaxation. Because the pseudopotentials for the common anion P are fit separately for GaP and InP, and thus non-identical, the P pseudopotential is taken as a weighted average in a combined system according to the number of Ga and In atoms on the four nearest-neighbor cation sites, in order to account for the difference in the local chemical environment.[13] The strain term is set to zero for the common anion P pseudopotentials. The pseudopotentials were obtained by fitting to experimentally determined or theoretically calculated electronic properties at their equilibrium conditions. These properties include energies, deformation potentials, effective masses at different critical points, and valence band offsets. The pseudopotentials can reproduce very well not only the binary band structures, but also the alloy band structure in the whole composition range with varying degree of order.[6, 14] A plane-wave basis is used to expand the electronic wavefunction, with a kinetic-energy cutoff of 7 Ry.

In the supercell calculation, if all the alloy eigenstates  $\{\psi_i\}$  need to be calculated to obtain the projection in Eq.(1), the calculation will be extremely expensive. In that case, one is either limited to analyze only a few special  $\mathbf{k}$  points ( $\Gamma$ , X, L) while using a large supercell (e.g.,  $\sim 30,000$  atoms) [10, 15] or forced to use smaller supercells (e.g.,  $\sim 4,000$  atoms or below) in order to get the full dispersion curves.[11] For instance, in our previous effort studying the band structure of  $\text{Ga}_x\text{In}_{1-x}\text{P}$  with  $x = 0.8$  using a 27648-atom supercell, even though the  $\Gamma$ -like state located at merely  $\sim 0.1$  eV above the band edge,  $A(\Gamma, E)$  spreads over a large spectrum range, thus, we had to calculate 200 states over a 230 meV spectral range for this large supercell.[15]

Even with the use of a very efficient computational technique – the folded spectrum method,[16] which allowed us to compute only the states close to a selected reference energy, the computation effort was still a major undertaking. Furthermore, we had to analyze the projections of all the 200 computed states in order to obtain the spectrum function  $A(\Gamma,E)$  in the spectral window of interest, which is also time consuming. Often, one cannot tell *in priori* where the peak of  $A(\mathbf{k},E)$  should appear and how many states to be calculated, if practically feasible to calculate the required number of states. In this work, we instead adopt the *generalized moment method* to calculate the spectral density function  $A(\mathbf{k}, E)$ . In this method, we first generate the VCA state  $\phi_n(\mathbf{k},\mathbf{r})$  within the primary cell, then extend it over the whole supercell. Note that this procedure works for any arbitrary  $\mathbf{k}$  point of the VCA BZ. If the  $\mathbf{k}$  point does not folded to the  $\Gamma$  point of the supercell, then a corresponding  $\mathbf{k}'$  point of the supercell cell BZ needs to be used. Assuming that the supercell Hamiltonian is  $H$  (which has been rescaled, so the minimum and maximum eigen energies are within  $[-1,1]$ ), we can generate the following Chebyshev wave functions via a recursion method:[17]

$$\phi_1 = H\phi_0, \quad (2)$$

$$\phi_j = 2H\phi_{j-1} - \phi_{j-2} = T_j(H)\phi_0. \quad (3)$$

Here  $\phi_0 = \phi_{nk}$ , and  $T_j(x)$  is the Chebyshev polynomial. Now, we can construct the  $j$ -th moment:

$$I_j = \langle \phi_0 | \phi_j \rangle = \langle \phi_{nk} | T_j(H) | \phi_{nk} \rangle = \sum_i \langle \phi_{nk} | \psi_i \rangle \langle \psi_i | T_j(H) | \phi_{nk} \rangle = \sum_i T_j(\epsilon_i) |\langle \phi_{nk} | \psi_i \rangle|^2. \quad (4)$$

Using the definition of Eq.(1), we have:

$$I_j = \int_{-1}^1 T_j(E) A_n(\mathbf{k}, E) dE. \quad (5)$$

This is the Chebyshev moment of the spectral function  $A_n(\mathbf{k},E)$ . With many of such moments (up to  $m$ -th order), we can reconstruct the spectral function  $A_n(\mathbf{k},E)$  as:

$$A_n(\mathbf{k}, E) = \frac{2}{\pi} (1 - E^2)^{-1/2} \sum_{j=0, m} T_j(E) I_j(1 + \delta_{j,0}). \quad (6)$$

The size of  $m$  determines the energy resolution of  $A_n(\mathbf{k}, E)$ . That  $m$  being a finite number leads to broadening each of the  $\delta$  function in  $A_n(\mathbf{k}, E)$  in Eq.(1) at  $E = E_i$  into a Gaussian function:

$$f(E') = e^{-\frac{(E' - E_i)^2}{\gamma^2}}, \quad (7)$$

where  $\gamma = 4 \frac{E_{\max} - E_{\min}}{m} \sqrt{1 - \left( \frac{E_{\max} - E'}{E_{\max} - E_{\min}} \right)^2}$ ,  $E_{\max}$  and  $E_{\min}$  are, respectively, the maximum and

minimum energy of the full spectrum that the alloy states span. In the language of the optical spectroscopy, the parameter  $\gamma$  determines the spectral resolution with a full width at half maximum (FWHM) given by  $2\sqrt{Ln2}\gamma$ , and  $m$  determines the step size given by  $\Delta = (E_{\max} - E_{\min})/m$ , typically about FWHM/6. To resolve the alloy states of a large supercell, a large  $m$  is desirable. For calculating  $A(\mathbf{k}, E)$ , we have used  $m = 200000$  (FWHM  $\approx 2.3$  meV) for 27000-atom supercells, and  $m = 300000$  (FWHM  $\approx 1.5$  meV) for the largest supercell of 259200 atoms, in an attempt to obtain  $A(\Gamma, E)$  for the conduction band band edge state at  $x = 0.5$ . For large  $m$ , Eq.(6) can be carried out via a fast Fourier transformation.[17] We note that in this moment method, the numerical error is not amplified through the recursive application of the Hamiltonian, rather a simple accumulation of those numerical errors from individual iterations. Even with the use of a large number of moments (e.g.,  $m = 300000$  or larger), because of the use of double precision, there is no problem for the numerical stability and accuracy.

In the configuration averaging approach,  $A_n(\mathbf{k}, E)$  can be obtained by averaging over many configurations:

$$\bar{A}_n(k, E) = \langle A_n(k, E) \rangle_{config}. \quad (8)$$

In this work, we use Eq.(1) for supercell size > 27000 atoms, Eq.(8) for 3456-atom supercells by averaging over 50 randomly generated configurations. In our previous work,[6, 10, 15] for convenience in studying the CuPt ordering, an orthorhombic supercell is adopted with three cell vectors  $\mathbf{a}_1$ ,  $\mathbf{a}_2$  and  $\mathbf{a}_3$  along the  $x' \sim [11\bar{2}]$ ,  $y' \sim [\bar{1}10]$  and  $z' \sim [111]$  direction of the zinc-blende (ZB) crystal, respectively. The supercell containing 27,648 atoms has  $a_1 = 12\sqrt{3/2}a$ ,  $a_2 = 12\sqrt{2}a$ , and  $a_3 = 8\sqrt{3}a$ , where  $a$  is the lattice constant of the alloy that is assumed to follow the Vegard's rule with  $a_{\text{GaP}} = 5.447 \text{ \AA}$  and  $a_{\text{InP}} = 5.8658 \text{ \AA}$ . Smaller supercells with a factor 2 and 4 reduction along all directions are also used with 3456 and 432 atoms, respectively. In this work, we also use two cubic supercells: with  $a_1 = a_2 = a_3 = 15a$  or 27000 atoms and  $a_1 = 144a$ , and  $a_2 = a_3 = 15a$  or 259200 atoms. A valence force field method is applied to relax all the atoms within the supercell to minimize the strain energy.[18] In the supercell, the total number of the cation atoms is enforced to satisfy the composition  $x$  and they randomly occupy the corresponding sub-lattice. Virtual crystal calculations are also performed, with the pseudopotential given by  $v(q, \boldsymbol{\epsilon}) = xv^{\text{GaP}}(q, \boldsymbol{\epsilon}) + (1-x)v^{\text{InP}}(q, \boldsymbol{\epsilon})$ .

### III. Results and Discussions

#### 1. Intrinsic broadening due to alloying

As pointed out in the introduction, for different purposes and convergence expectations, different supercell sizes are required. Figure 1 shows the histogram plots for the bandgaps of the disordered  $\text{Ga}_{0.5}\text{In}_{0.5}\text{P}$  calculated using two supercell sizes, Fig.1(a) for the 432-atom supercell and Fig.1(b) for the 3456-atom supercell, each with 100 configurations. When the size increases from 432 atoms to 3456 atoms, not only the energy fluctuation decreases but also the average bandgap increases from 1.971 to 1.980 eV. The latter is very close to the result of one even larger supercell – 27648-atom, 1.979 eV.[6] Therefore, for the purpose of determining the

bandgap, the configuration averaging using the 3456-atom supercell is considered adequate. The full width at half maximum for the histogram plot of Fig.1(b) is  $\Delta_{cf} = 8.3$  meV,[19] which turn out to be similar to the excitonic linewidth  $\Delta_{ex} = 8$  meV for this alloy measured by the excitonic emission at low temperature.[20] It is reasonable to expect there is some connection between  $\Delta_{cf}$  and  $\Delta_{ex}$ , but we will leave this aspect for a future work. We only wish to point out here that  $\Delta_{cf}$  is rather different from the most popularly considered mechanism for the excitonic linewidth in a semiconductor alloy where the width is caused by the energy variation as a result of composition fluctuation within the “exciton volume” from one area to the other.[21] In our case, the average composition within the supercell is kept the same and the bandgap fluctuation is purely due to the arrangement of the atoms in the microscopic scale. In this work, we instead focus on the difference between  $\Delta_{cf}$  and the width of  $A(\Gamma,E)$ , the intrinsic broadening of the band edge alloy state, to be investigated below. To determine  $A(\Gamma,E)$ , not only the 3456-atom cell even the 27648-atom cell is far from sufficient.

Figure 2(a) shows  $A(\Gamma,E)$  for the conduction band minimum (CBM) calculated from one 3456-atom and one-27648-atom supercell, with  $A(\Gamma,CBM) = 0.867$  and  $0.864$ , respectively. Although it appears the CBM has a large component of the VCA  $\Gamma$  state, the results in fact represent the integrated VCA  $\Gamma$  components for all the alloy states within a spectral range of  $\sim 20$  or  $\sim 145$  meV, respectively, for the two supercell sizes. The range is roughly given by the energy separation between the band edge state and the nearest alloy state, as shown in Fig.2(a), thus, depending on the size of the supercell. In order to be able to investigate the intrinsic broadening of the alloy CBM state in this direct bandgap semiconductor, one would need to use even much larger supercell. If we would like to examine the coupling of the CBM to other alloy states lying within 1-2 meV of the CBM, the supercell dimension needs to be increased by roughly a factor

of 10, assuming the effective mass is in the order of  $m^* \sim 0.1$ . Therefore, we have tested two supercell sizes based on an 8-atom cubic cell with size  $15 \times 15 \times 15$  and  $144 \times 15 \times 15$ , using the much more efficient GMM. The results are shown in Fig.2(b). Note that in Fig.2(b), each peak represents one alloy state that has non-zero projection to the VCA CB  $\Gamma$  state, and is broadened by the lineshape function of Eq.(7). Again, the 27000-atom cell does not reveal any alloy state within the vicinity of the CBM, but the 259200-atom cell does give rise to four peaks within 25 meV. These four peaks can be understood as resulting from alloying induced coupling among the folded VCA states (intra-valley coupling) with  $k_x = j/144$  ( $2\pi/a$ ) and  $k_y = k_z = 0$ , where  $a$  is the lattice constant of the VCA crystal and  $j = 0,1,2,3$ . Because the coupling to states of higher  $j$  values diminishes quickly, we cannot obtain the dispersion curve of the alloy states far away from the CBM. Nevertheless, we can have an estimate for the  $\Gamma$  valley effective mass of the alloy as  $m(\Gamma, x = 0.5) = 0.0994 m_0$ ,  $m_0$  is the free electron mass, which is very close to the VCA effective mass  $m_{\text{VCA}}(\Gamma, x = 0.5) = 0.103 m_0$ . Alternatively, one can use a smaller supercell, the 27000-atom one for instance, to obtain the dispersion curve of the alloy by applying the GMM using VCA states at different  $\mathbf{k}$  points, which will be illustrated later for the effort to explore the global electronic structure of the alloy.

The result of Fig.2(b) allows us to estimate the intrinsic broadening at the CBM due to the intra-valley scattering within the  $\Gamma$  valley, which can be viewed as *inelastic* scatterings among the VCA states. The standard CPA theory for the alloy predicts in the weak scattering limit a Lorentzian lineshape for all the  $\mathbf{k}$  states.[1] However, because the coupling to the states below the VCA (e.g., the valence band states) is very weak, the  $A(\Gamma, E)$  spectrum is expected to be non-Lorentzian and asymmetric. For  $E < 0$ , we can assume  $A(\Gamma, E) = 0$  for most practical purposes. For  $E \geq 0$ , a continuous spectrum of  $A(\Gamma, E)$ , equivalent to extrapolating to an infinite size

supercell, can be obtained by finding the envelop of the  $j = 0 - 3$  peaks of Fig.2(b). Because of the missing of the folded states along the  $k_y$  and  $k_z$  direction, the intensities of the  $j = 1 - 3$  should be approximately 3 times as strong as those of Fig.2(b), if all three directions are included. We find that the spectral function can be described by a stretched exponential function:

$$A(\Gamma, E) = A_0 e^{-\alpha E^\beta}, \quad (9)$$

with  $\alpha = 1.162$ ,  $\beta = 0.4842$ , and  $A_0 = 0.5663 \text{ (meV)}^{-1}$ , where  $E$  is in meV. The integrated intensity is 0.857 for the range of 25 meV measured from the CBM. The width at the half maximum is  $\Delta_{\text{CB-}\Gamma} = 0.34 \text{ meV}$ , which measures the alloying induced intrinsic broadening of the band edge state, as a result of the intra-valley coupling. Note that, in the standard VCA based scattering treatment, at the  $\Gamma$  point, there would be no alloy scattering. However, in our case, the spectral width at the  $\Gamma$  point is finite, although small, which will result in a finite quasi-particle life time at the  $\Gamma$  point. Clearly  $\Delta_{\text{CB-}\Gamma}$  is very different from  $\Delta_{\text{cf}}$  in origin. We note that  $\Delta_{\text{CB-}\Gamma}$  corresponds to an infinitely large alloy, whereas the energy fluctuation in the length scale of other co-existing physical mechanisms, such as excitonic effect or electron-phonon scattering, are more relevant to the excitonic linewidth or transport in reality.

We may use the formalism derived from the single site scattering theory to approximately describe the decay of the  $\Gamma$ -like state:[1]

$$I(\mathbf{k}, t) = \left| \int_0^\infty dE e^{-\frac{iEt}{\hbar}} A(\mathbf{k}, E) \right|^2. \quad (10)$$

Contrary to the ideal case,[1] the time dependence is non-exponential, as shown in Fig.3, but the  $1/e$  decay time, which may be considered as scattering time  $\tau_s(\mathbf{k})$ , is found to be 0.21 ps for the  $\Gamma$  point. One could understand this decay time as the lifetime of a VCA state that is viewed as a wave packet constructed by the alloy eigenstates within a small energy spread. After an average

under electron thermal distribution (at different  $\mathbf{k}$  point near  $\Gamma$ ),  $\tau_s$  is often used to estimate the electron mobility  $\mu_e$ , which is simply  $e\tau_s/m^*$  in Drude model. [2] Note that this scattering time, which is also known as “quasi-particle lifetime”, [1] is quite different from the real carrier lifetime due to either radiative recombination with a hole or relaxation to defect or impurity states below the CB band edge. Here, the electron is scattered into a different state (e.g., in Boltzmann equation), while for the real carrier lifetime, the carrier is eliminated. In fact, the real carrier lifetime is much longer in a high quality alloy sample. For instance, the electron lifetime measured by radiative decay of the photoluminescence at low temperature (e.g. 1.5 K) is in the order of 200 ps for  $\text{Ga}_{0.5}\text{In}_{0.5}\text{P}$  alloy. [22]

## 2. Alloy states far away from the band edge

Next, we apply the GMM to calculate the electronic “dispersion” relations along two high symmetry lines  $\Gamma$ -X and  $\Gamma$ -L of the BZ. Figure 4 shows the alloy “dispersion” curves for  $\text{Ga}_{0.5}\text{In}_{0.5}\text{P}$ , comparing with the results of VCA with and without strain effect (i.e., the strain parameter  $a_s \neq 0$  or  $a_s = 0$  in the atomic pseudopotential). The three curves appear to have a similar shape, but the difference between the alloy and VCA can be more than 0.5 eV and is non-constant throughout the BZ. The energy state on the alloy curve for a given  $\mathbf{k}$  is determined from the peak position of the spectral function  $A(\mathbf{k},E)$ , using the 27000-atom supercell.

The spectral function  $A(\mathbf{k},E)$  at the band edge of a direct bandgap alloy has been shown to be very sharp, as illustrated in Fig.2(b). Figure 5 shows how  $A(\mathbf{k},E)$  evolves along  $\Gamma$ -X and  $\Gamma$ -L lines, computed by using the 27000-atom supercell and  $m = 200,000$ . For those  $\mathbf{k}$  points very close to the  $\Gamma$  point, for instance  $\mathbf{k} = 0.05 \mathbf{k}_L$ , the spectra show a single peak and its linewidth is dictated by the “spectral resolution” related to  $\gamma$  in Eq.(3). The peak position gives correctly the

energy level on the dispersion curve, though the peak height is affected by the inadequate supercell size. For those  $\mathbf{k}$  points slightly further away from the  $\Gamma$  point, for instance  $\mathbf{k} = 0.10 \mathbf{k}_L$ , the coupling to the adjacent  $\mathbf{k}$  points are apparent, even in the linear plot, indicating enhancement in the “intra-valley scattering”. In general, the further away from the band edge is a VCA state, the broader the distribution of the VCA state is in the alloy energy spectrum. For the  $\mathbf{k}$  points close to the highest energy point of the dispersion curve, for instance  $\mathbf{k} = 0.40 \mathbf{k}_L$ , the  $\mathbf{k}$  VCA state is found to disperse into alloy states in a spectral range more than 0.5 eV. The peak height has reduced from  $A_0 \sim 566 \text{ eV}^{-1}$  at  $\Gamma$  to  $A_0 < 5 \text{ eV}^{-1}$  at  $\mathbf{k} = 0.4 \mathbf{k}_L$ , which means that the alloy state with the largest  $\mathbf{k}$  component contains only less than 1% of the  $\mathbf{k}$  component. In a spectroscopy measurement, such alloy effect will reflect in the reduction in the signal intensity and linewidth broadening. The detail depends on the type of measurement, which is beyond the scope of this work.

It is of particular interest to examine  $A(\mathbf{k}, E)$  at a critical point, because it provides more clear picture for the inter-valley and intra-valley scattering than at a general  $\mathbf{k}$  point. Figure 6 shows the spectral functions for the X and L point, calculated for all the VCA states belonging to different degenerate valleys,  $X_1 - X_3$  and  $L_1 - L_4$ . For the L-like alloy state at around 0.22 eV above the band edge, as shown in Fig.6(a), the alloying effect is dominant by the inter-valley scattering, as indicated by the four separated peaks spreading out in a range of  $\sim 10 \text{ meV}$  and derived from the four L valleys. For the X-like state at around 0.37 eV above the band edge, as shown in Fig.6(b), there are apparently more alloy states involved, spreading out in a range greater than 40 meV, and one could not unambiguously identify three X-like peaks. The larger number of peaks for the X-like state, contrasting to the L-like state, could be understood as more significant intra-valley scattering, because on the one hand, the X-like state is further away from

the band edge; on the other hand, in VCA, as shown in Fig.4(a), the X point is not the local minimum along the  $\Gamma$ -X line, which leads to large number of states that are energetically degenerate with the X point with their  $\mathbf{k}$  vectors close to that of the X point.

From the alloy dispersion curves shown in Fig.4, we can calculate the effective masses near the  $\Gamma$  point along both the  $\Gamma$ -X and  $\Gamma$ -L lines, and even examine the possible anisotropy and non-parabolicity of the dispersion, which would not be possible before. As references, we find the VCA effective masses to be 0.103 and 0.106, respectively, for the [0,0,1] and [1,1,1] directions, and the non-parabolicity starts at around 5%  $k_X$  or  $k_L$ . The masses for the alloy are practically the same as the VCA results:  $0.103 \pm 0.001$  and  $0.106 \pm 0.001$  for the [0,0,1] and [1,1,1] directions, which means that the alloy band dispersion near the BZ center is very close to that of the VCA and nearly isotropy, but with a rigid shift in energy from the VCA dispersion curves. Also note that the [0,0,1] effective mass agrees well with that estimated from the spectral function of the  $\Gamma$  point. We expect that the proper supercell size for computing the effective mass should be the similar to that for the bandgap, because the states at different  $\mathbf{k}$  points experience the same alloy configuration, thus the energy difference between close by  $\mathbf{k}$  points is not as sensitive to the alloy configuration fluctuation as the bandgap itself. . In general, the higher lying states tend to require larger supercells for achieving the same convergence as the band edge state. However, one might not need to achieve the same convergence in practice, because the higher lying alloy states tend to suffer more alloy broadening, as is evidenced in optical spectroscopy measurements.

### **3. Indirect band-gap alloy**

Above we have discussed the intra- and inter-valley scatterings when the conduction band X- and L-like states are away from the lowest band edge. We now investigate the alloying effects for the composition region that resembles an indirect bandgap semiconductor, using  $x_{\text{Ga}} = 0.8$  as

an example and focusing on the conduction band. The band structures have been calculated in three ways: (1) averaging over 50 configurations of the 3456-atom supercell, (2) average of two 27648-atom supercells, and (3) one 27000-atom supercell. The lowest CB state energies are found to differ by less than 4 meV (-3.7225, -3.7187, -3.7205 eV) among the three results, which suggests that they are adequately converged. The first three CB states are found all having  $A(k_X, E_i) \sim 0.9$  (summing over the three X valleys). Thus, they can be understood as split  $X_1$ -like states due to the inter-valley scattering. However, the splittings are significantly larger for the 3456-atom supercell than for the larger supercells, for instance 7.1 meV vs. 1.7 meV between the first two states, as shown in Fig.7. This is another example showing the non-equivalency of the two approaches. One can envision that the inter-valley splitting will diminish if the supercell size is increased further, indicating a macroscopic alloy should have in average the zinc-blende symmetry. However, the observed supercell size dependence of the inter-valley splitting suggests that the impact of the alloy scattering depends on the spatial extension of a physical property that is used for probing. For instance, an excitonic state in a semiconductor alloy could be viewed as a probe for the alloy fluctuation with an approximate probe size of the Bohr radius of the exciton that is about the size of the 27000 atom supercell;[21] whereas a defect or an isoelectronic impurity state could be sensitive to the alloy fluctuation in a substantially small scale, for instance, a N impurity in  $Ga_xIn_{1-x}P$  can sense the Ga and In coordination change in its first nearest neighbors.[23]

The additional peaks (calculated by GMM) shown in Fig.7 at the higher energy side of the  $X_1$ -like alloy states are alloy states corresponding to the VCA states with  $\mathbf{k}$  vectors such as  $(\pm 14/15, 0, 0)$ , coupling with the  $X_1$  states through the intra-valley scattering. There are total six of them in this group (although not all of them are resolved in this plot), spanning an interval of

about 9 meV. Compared to Fig.6(b) where the  $X_1$ -like states are away from the band edge in the direct composition region, now in the case of indirect composition region, for the alloys states near the CBM the alloy scattering is significantly weaker (i.e., only those energetically close by  $\mathbf{k}$  points are involved). However, the  $A(k_x, E)$  spectra shown in Fig.7 indicate that the intra-valley scattering is in fact comparable with the inter-valley scattering.

## 4. Discussions

### 4.1 Alloy scattering theories within the framework of VCA

For an alloy  $A_xB_{1-x}C$ , the perturbation potential is  $\Delta V = V - V_0$ , where  $V$  and  $V_0$  are respectively the total alloy and VCA potential. Assumed no lattice relaxation, the perturbation matrix element is often given as

$$|\langle \mathbf{k}' | \Delta V | \mathbf{k} \rangle|^2 \approx Nx(1-x) |\delta V_{AB}(\mathbf{k}', \mathbf{k})|^2 \quad (11)$$

where  $\delta V_{AB} = V_A - V_B$  ( $V_A$  and  $V_B$  are atomic potentials for atom A and B), and the matrix element  $\delta V_{AB}(\mathbf{k}', \mathbf{k}) = \langle \mathbf{k}' | \delta V_{AB} | \mathbf{k} \rangle$ . Note that Eq.(11) is valid only for  $\mathbf{k}' \neq \mathbf{k}$ , under the assumption that the two-body terms involving two atoms on different sites (A-A, B-B, and A-B) are all negligible.[24] The diagonal matrix element  $\langle \mathbf{k} | \Delta V | \mathbf{k} \rangle = 0$  (although  $\langle \mathbf{k} | \delta V_{AB} | \mathbf{k} \rangle \neq 0$ ), because  $\langle \mathbf{k} | \Delta V | \mathbf{k} \rangle$  is simply the average of the potential fluctuation with respect to  $V_0$ , which perhaps provides the base for the common use of an elastic scattering theory.

We first consider the theory for no lattice relaxation. In the context of the elastic scattering theory (within the first Born approximation), the matrix element square given by Eq.(11) is often related to the alloy scattering rate  $1/\tau$  or scattering induced spectral broadening  $\Delta$ .[8] Using this approach, because of the constrain to elastic scattering, the linewidth  $\Delta$  is zero for the non-degenerate  $\Gamma$  valley, and only the inter-valley scatterings among degenerate valleys can yield

non-zero  $\Delta$ . [8] Apparently, in this approach, the effect of inelastic intra-valley scatterings is not present. However, it has been shown above that the inelastic intra-valley scattering is actually quite significant for an indirect bandgap alloy (Fig.7), and in a direct bandgap alloy it is the intra-valley scattering that generates the finite linewidth  $\Delta_{CB-\Gamma}$  for the  $\Gamma$  point (Fig.2). We would like to mention that in the literature for the weak potential fluctuation the spectral width of an alloy state is also given as  $\Delta(\mathbf{k}) = \pi x(1-x) |\langle \mathbf{k} | \delta V_{AB} | \mathbf{k} \rangle|^2 \rho(E)$ , from the Green's function theory, where  $\rho(E)$  is the density of states. [2, 8] Because  $\rho(E)$  approaches zero for the  $\Gamma$  point,  $\Delta(\Gamma)$  approaches zero too, [2] although one could obtain a non-zero  $\Delta(\Gamma)$  by introducing an imaginary part into  $E$  (termed as “band tailing”). [8] Note that  $\Delta_{CB-\Gamma}$  and  $\Delta(\Gamma)$  are very different in nature. The former is the result of inelastic scattering, and the latter of elastic scattering. Also note that in the alloy scattering discussion, the “elastic” and “inelastic” scatterings are not actually associated with the energy change of an alloy state, but the coupling among the VCA states, either degenerate or non-degenerate, induced by the potential fluctuation.

With lattice relaxation, the diagonal term  $\langle \mathbf{k} | \Delta V | \mathbf{k} \rangle$  becomes non-zero, typically a few hundred meV. [15] For the indirect bandgap case considered above with  $x = 0.8$ , we have found that the diagonal and off-diagonal matrix elements among the  $X_1$  states in fact have the same order of magnitude,  $|\langle \mathbf{k}' | \Delta V | \mathbf{k} \rangle| \sim 0.1 - 0.2$  eV. These results indicate that the conventional approach inappropriately neglects the inelastic scattering that is present even for the alloy without lattice relaxation, and the treatment becomes more problematic when the lattice relaxation is present.

#### **4.2 An alloy transport theory not relying on a reference system?**

Until now, we have followed the common practice in the literature to discuss the alloy scattering in the framework of the VCA. We next point out a more fundamental issue of the

electronic transport theory within the framework of the VCA. Because of the use of the virtual crystal structure as the reference, any deviation from the VCA will lead to non-zero scatterings among the VCA states, thus, implying a reduction in carrier mobility. Obviously, an extreme example of non-random distribution would be a structure with long-range order, for instance, a semiconductor superlattice.[25] For this case, it is not really meaningful to discuss the “alloy scattering”. A more subtle situation is the well studied partial long-range ordering that has often been observed in III-V semiconductor alloys,  $\text{Ga}_x\text{In}_{1-x}\text{P}$  in particular.[26] To some extent, it is still useful to use the VCA states to characterize the alloy state in a partially ordered structure.[15] However, it would be a non-trivial task to separate the effects of long-range ordering and alloy fluctuation, if one attempts to calculate the scattering rate and further the carrier mobility, using the scattering matrix element based on the VCA. For a superlattice structure with imperfections such as atom inter-diffusion, it would be more nature to use the ideal superlattice as the reference rather than the VCA. There is apparently some arbitrariness and non-uniqueness using either reference for evaluating the transport property. And there is a disparity between the electronic transport and the electromagnetic wave propagation. For the latter, the energy propagation can be calculated without relying on any reference system whether or not the medium is disordered or ordered. Therefore, there is a need to develop a theory that is independent of any reference system for the electronic transport in a non-periodic structure. The large supercell direct calculation performed in this work can play an important role in this regard.

#### **IV. Summary**

We have compared two supposedly equivalent approaches for describing (semiconductor) alloys: (1) configuration averaging over many “appropriate size” supercells versus using one “very large size” representative supercell. We have found that the specific size for either

“appropriate size” or “very large size” depends on the problem of interest. For certain properties, such as bandgap, the two approaches are indeed capable of giving equivalent results. However, for others, such as properties related to alloy statistics, they are not at all equivalent.

We have applied two techniques to analyze the global electronic structure of a semiconductor alloy system  $\text{Ga}_x\text{In}_{1-x}\text{P}$ . A *generalized moment method* (GMM) is used to directly calculate the spectral function  $A(\mathbf{k},E)$  of the alloy in the whole spectral range without having to explicitly calculate any alloy state, and can deal with a very large system with  $> 100,000$  atoms (up to a few million atoms). The results can yield bandgap, dispersion curves (by projecting alloy states into the VCA states), and intrinsic alloy broadening that naturally includes both elastic and inelastic scatterings. For a representative composition  $x = 0.5$  in the direct bandgap region, we find that the intrinsic broadening for the  $\Gamma$ -like alloy state, the width of the spectral function  $A(\Gamma,E)$ , is non-Lorentzian and asymmetric, with an estimated width  $\Delta_{\text{CB-}\Gamma} = 0.34$  meV and a corresponding “quasi-particle” lifetime  $\tau_s = 0.21$  ps. However, the width of the spectral function in the middle point of the BZ could be more than 0.5 eV. The alloy dispersion curve near the  $\Gamma$  point is nearly a rigidly shifted one of the VCA with  $m^* = 0.10 m_0$ . We point out that the intrinsic alloy broadening  $\Delta_{\text{CB-}\Gamma}$  is fundamentally different from the energy fluctuation  $\Delta_{\text{cf}}$  obtained by computing many “appropriate size” alloy configurations and  $\Delta(\Gamma)$  discussed in the literature. For a representative composition  $x = 0.8$  in the indirect bandgap region, we find that the “inelastic” intra-valley scattering, typically neglected in the conventional alloy transport theory, is in fact in the same order of magnitude as the “elastic” inter-valley scattering; and the inter-valley splittings obtained from configuration averaging over the “appropriate size” supercells ( $\sim 3500$  atoms) is substantially smaller than those from individual large supercells ( $\sim 27000$  atoms). It is satisfying to know that modern computation techniques like the one

demonstrated here can directly calculate a system with the size comparable to the extent of a physical interaction of interest (e.g., exciton). This capability can open the way for detailed comparisons between theory and experiment regarding the experimentally observed statistical fluctuations in an alloy system.

We further discuss the need to develop an alloy transport theory that does not rely on the use of a reference system, for instance, the popular virtual-crystal. We argue that the large scale supercell calculation can play an important role in such development.

**Acknowledgements** This work is supported partially by Charlotte Research Institute at UNC-Charlotte, and DOE-OS-BES-MSED under contract DE-AC02-05CH11231 to LBNL. The work used the computational resources of NERSC at LBNL.

## References

- [1] H. Ehrenreich, and L. M. Schwartz, *The electronic Structure of Alloys* (Academic Press, New York, 1976), Vol. 31.
- [2] A. B. Chen, and A. Sher, *Semiconductor alloys, physics and materials engineering* (Plenum, New York, 1995).
- [3] A. Zunger *et al.*, Physical Review Letters **65**, 353 (1990).
- [4] L. W. Wang *et al.*, Physical Review Letters **80**, 4725 (1998).
- [5] P. Ernst *et al.*, Applied Physics Letters **67**, 2347 (1995).
- [6] Y. Zhang, A. Mascarenhas, and L. W. Wang, Physical Review B **63**, R201312 (2001).
- [7] R. J. Lempert, K. C. Hass, and H. Ehrenreich, Physical Review B **36**, 1111 (1987).
- [8] C. H. Grein, S. Zollner, and M. Cardona, Physical Review B **44**, 12761 (1991).
- [9] S. Krishnamurthy, A. Sher, and A. B. Chen, Physical Review B **33**, 1026 (1986).
- [10] Y. Zhang, A. Mascarenhas, and L.-W. Wang, Physical Review Letters **101**, 036403 (2008).
- [11] V. Popescu, and A. Zunger, Physical Review Letters **104**, 236403.
- [12] K. C. Hass, H. Ehrenreich, and B. Velický, Physical Review B **27**, 1088 (1983).
- [13] L.-W. Wang, J. Kim, and A. Zunger, Physical Review B **59**, 5678 (1999).
- [14] M. A. Steiner *et al.*, Journal of Applied Physics **106**, 063525 (2009).
- [15] Y. Zhang, A. Mascarenhas, and L.-W. Wang, Physical Review B **78**, 235202 (2008).
- [16] L. W. Wang, and A. Zunger, Journal of Chemical Physics **100**, 2394 (1994).
- [17] L.-W. Wang, Physical Review B **49**, 10154 (1994).
- [18] P. Keating, Physical Review **145**, 637 (1966).
- [19] Y. Zhang, A. Mascarenhas, and L. W. Wang, Physical Review B **64**, 125207 (2001).

- [20] Y. Zhang *et al.*, Physical Review B **61**, 9910 (2000).
- [21] K. K. Bajaj, Materials Science & Engineering **R34**, 59 (2001).
- [22] P. Ernst *et al.*, Phys. Status Solidi B **193**, 213 (1996).
- [23] H. Mariette *et al.*, Physical Review B **31**, 5217 (1985).
- [24] G. L. Hall, Physical Review **116**, 604 (1959).
- [25] L. Esaki, and R. Tsu, IBM Res. Dev. **14**, 61 (1970).
- [26] A. Mascarenhas, and Y. Zhang, in *Spontaneous Ordering in Semiconductor Alloys*, edited by A. Mascarenhas (Kluwer Academic/Plenum Publishers, New York, 2002), p. 283.

## Figure captions

Fig. 1. (color online) Histogram plots of the energy distributions of the band gap ( $E_g$ ) for  $\text{Ga}_{0.5}\text{In}_{0.5}\text{P}$  alloy, calculated using two super-cell sizes.

Fig.2. (color online) Spectral functions  $A(k=0, E)$  near the conduction band minimum for  $\text{Ga}_{0.5}\text{In}_{0.5}\text{P}$  alloy: (a) obtained by projecting the alloy wavefunctions to the VCA state at  $k=0$ ; (b) calculated by applying the generalized moment method with the  $k=0$  VCA state as the reference.

Fig.3. The time decay of the  $\Gamma$ -like alloy state in  $\text{Ga}_{0.5}\text{In}_{0.5}\text{P}$  alloy.

Fig.4. (color online) The effective dispersion curves of  $\text{Ga}_{0.5}\text{In}_{0.5}\text{P}$  alloy along the  $\Gamma$ -X and  $\Gamma$ -L directions, calculated using the generalized moment method (calculated discrete points interpolated by solid lines). The dashed and dotted curves are the results of VCA with and without the inclusion of the strain term in the atomic pseudopotentials.

Fig.5. (color online) Spectral functions at various  $\mathbf{k}$  points in the VCA Brillouin zone calculated for  $\text{Ga}_{0.5}\text{In}_{0.5}\text{P}$  alloy: (a) for the (111) direction, (b) for the (100) direction. The energy reference is the conduction band minimum. A 27000-atom supercell is used.

Fig.6. (color online) Spectral functions for  $\text{Ga}_{0.5}\text{In}_{0.5}\text{P}$ , calculated using the generalized moment method, corresponding to (a) four VCA  $L_1$  states, showing the coupling due to the inter-valley scatterings; and (b) three VCA  $X_1$  states, showing more extensive decompositions of the VCA  $X_1$  states among the alloy states centered around -3.66 eV. The positions of the  $X_1$ -like alloy

states are estimated by performing weighted averaging over the alloy states. A 27000-atom supercell is used.

Fig.7. (color online) The effect of inter- and intra-valley scatterings in  $\text{Ga}_{0.8}\text{In}_{0.2}\text{P}$  alloy. Left axis (discrete data points): the total X components of the three  $X_1$ -like alloy states for 50 random configurations of the 3456-atom supercell. The solid vertical lines indicate their average energies, and the splitting are caused by the inter-valley scatterings. Right axis (continuous curves): spectral functions for the three VCA  $X_1$  states, calculated using the generalized moment method and the 27000-atom supercell. The dashed vertical lines indicate the positions of the three  $X_1$ -like alloy states in the alloy with the splitting caused by the inter-valley scatterings. Additional peaks on the higher energy side are produced by the intra-valley scatterings.

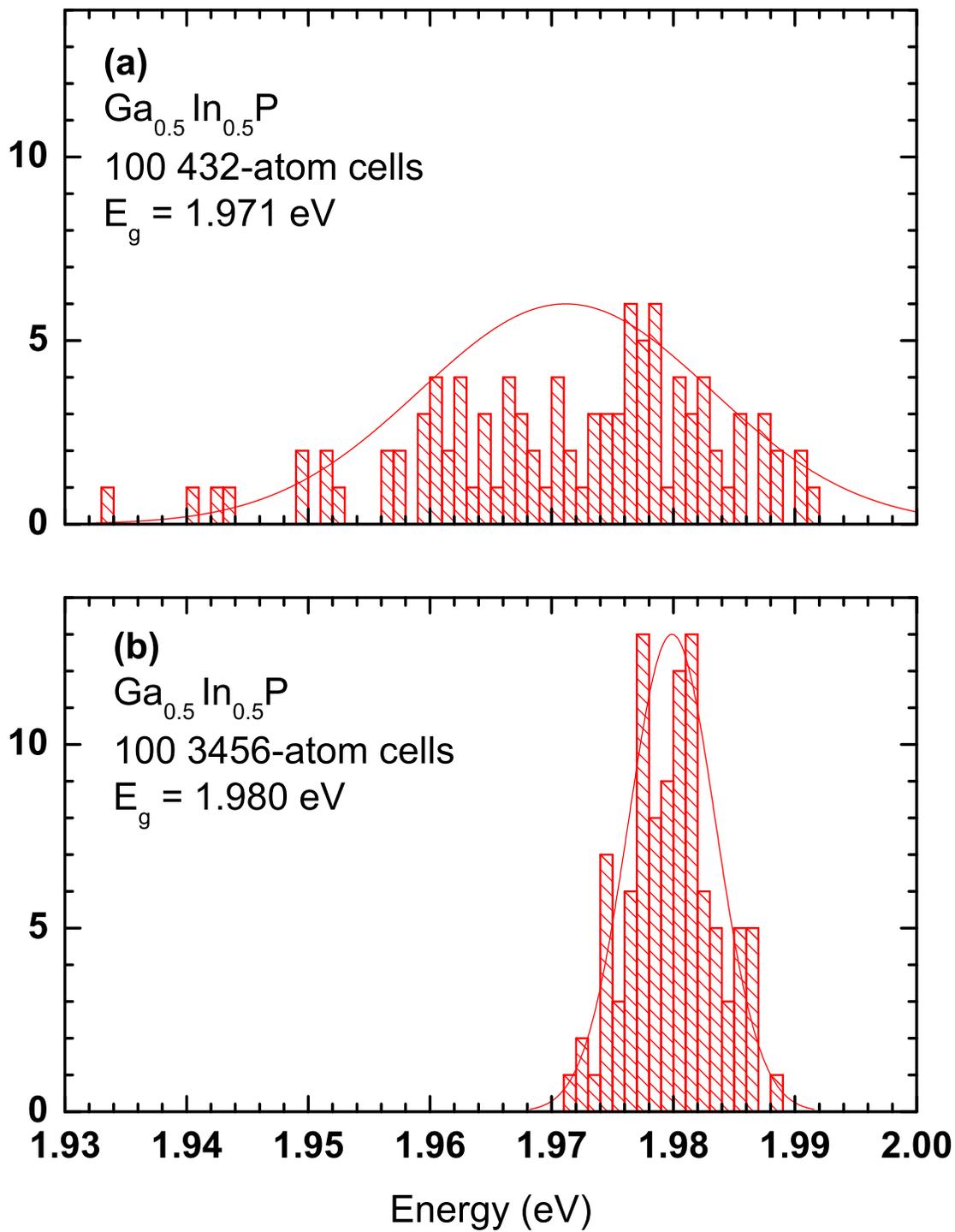


Figure 1

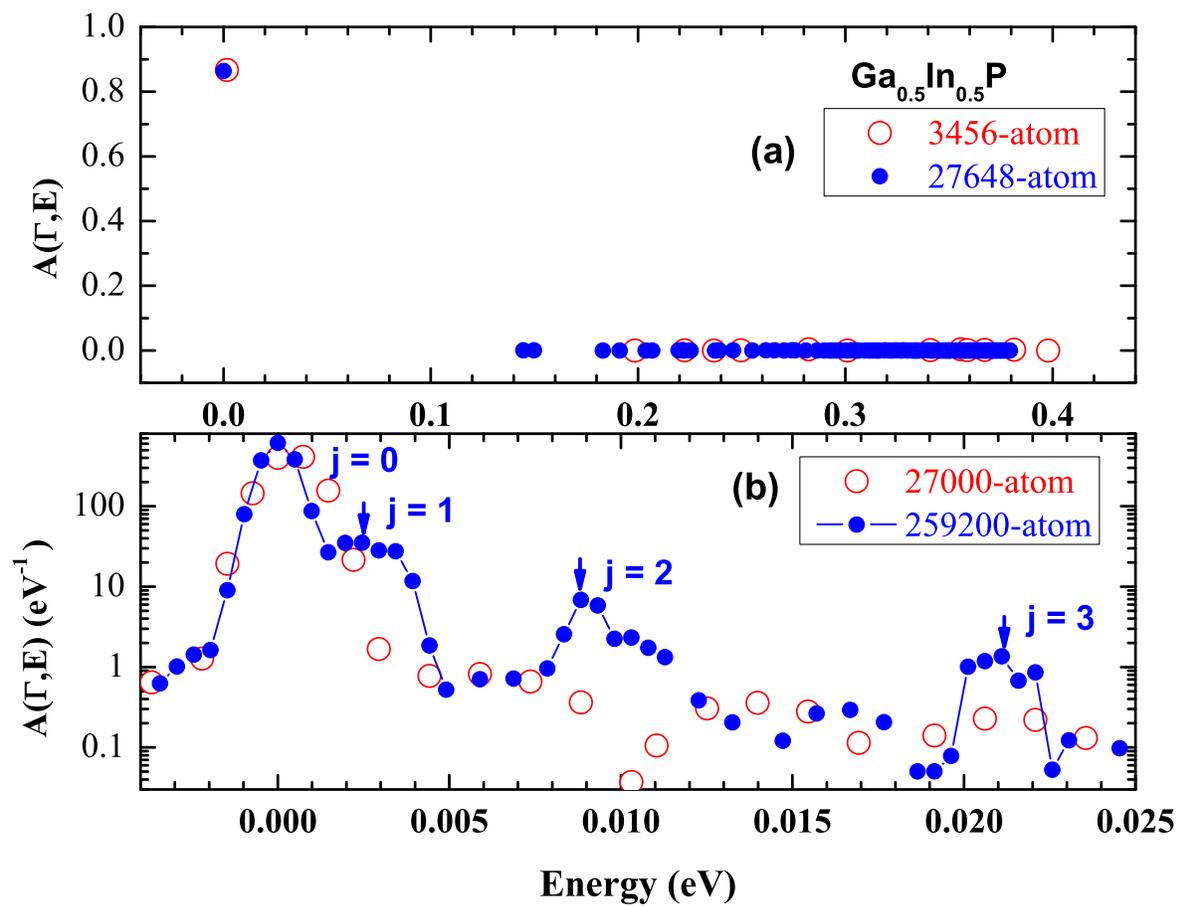


Figure 2

BB11500

03MAR2011

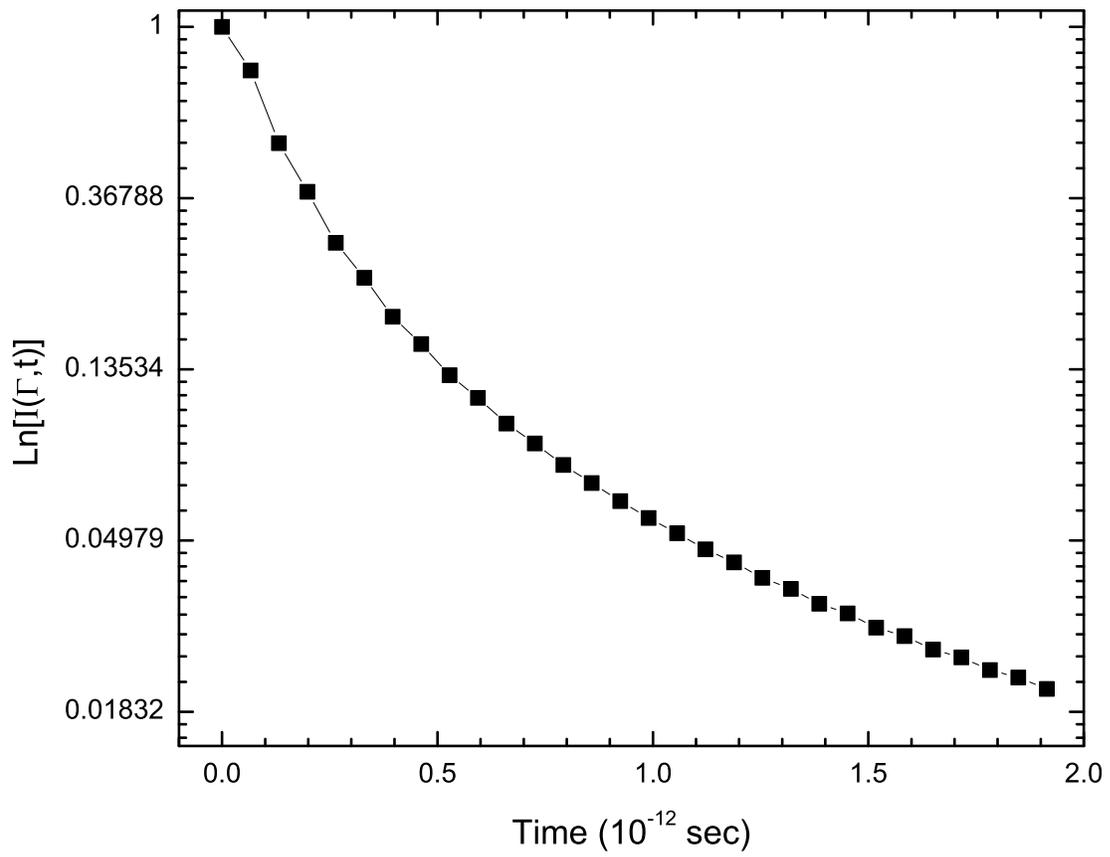


Figure 3 BB11500 03MAR2011

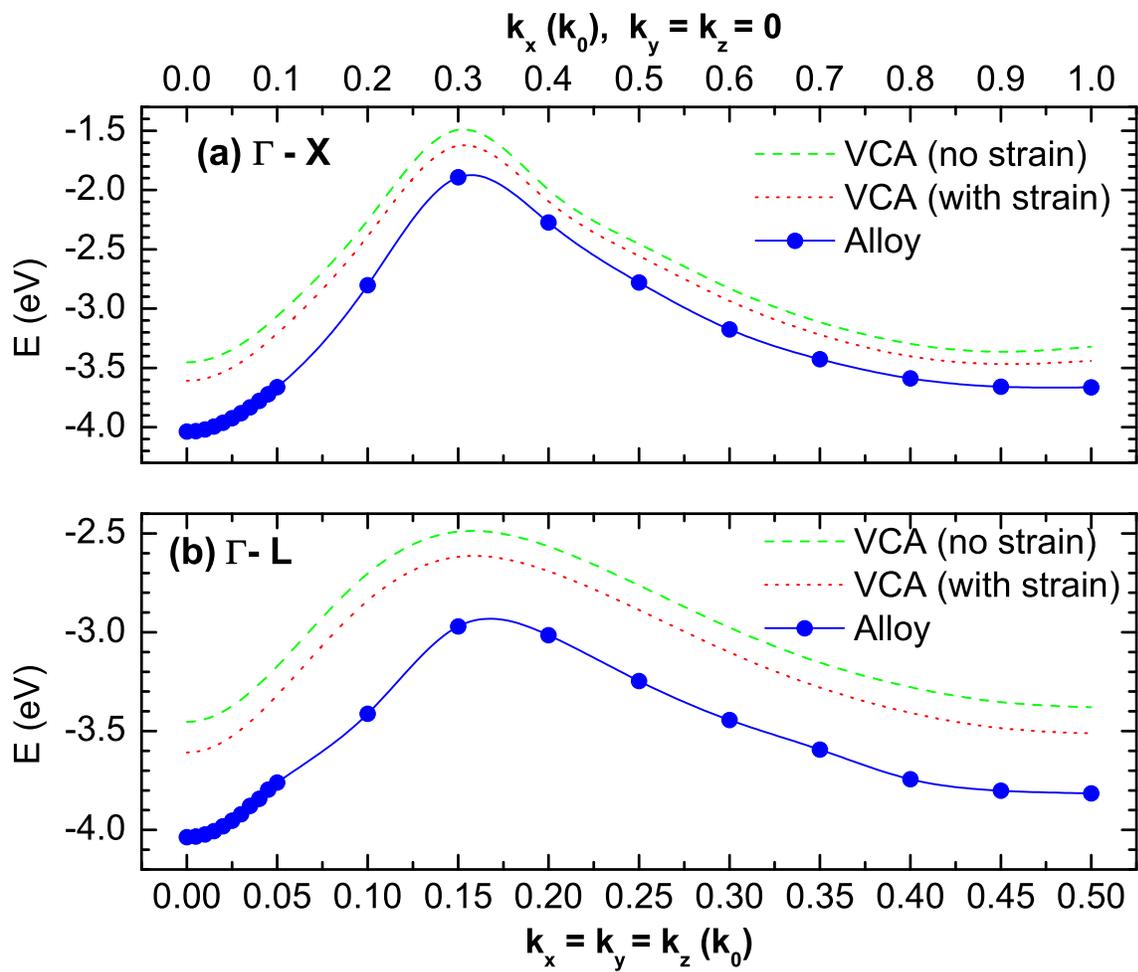


Figure 4

BB11500

03MAR2011

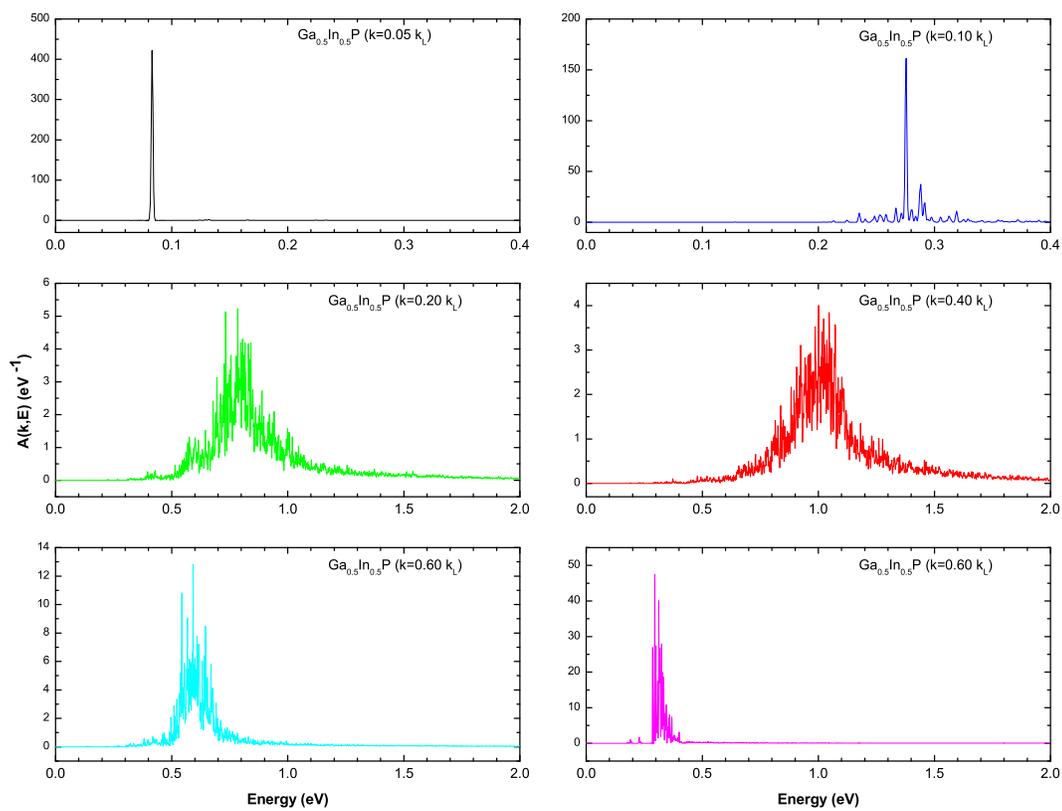


Figure 5a

BB11500

03MAR2011

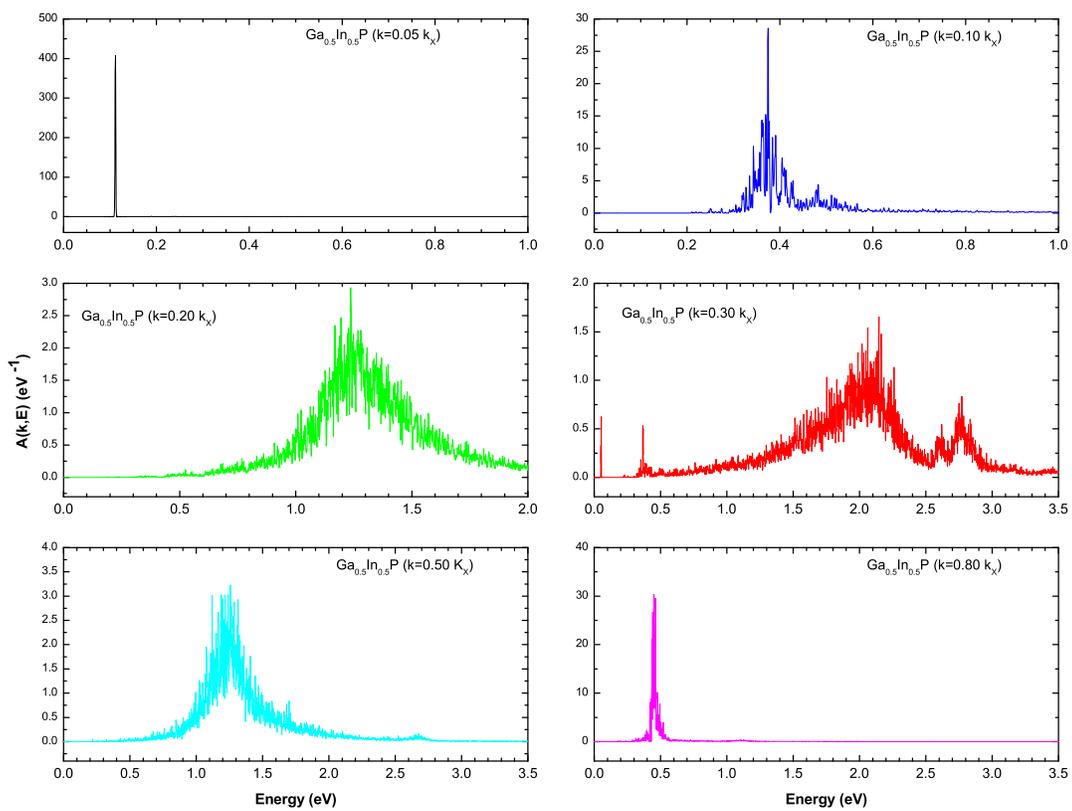


Figure 5b

BB11500 03MAR2011

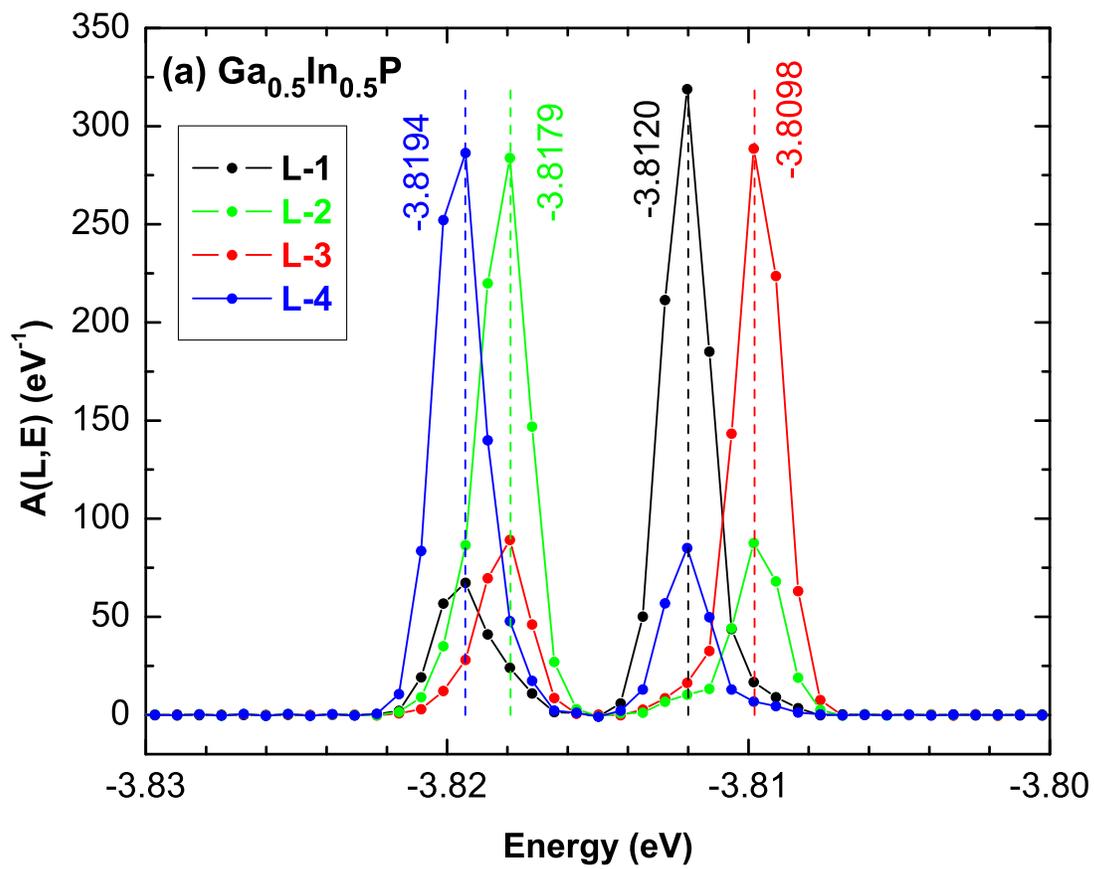


Figure 6a

BB11500

03MAR2011

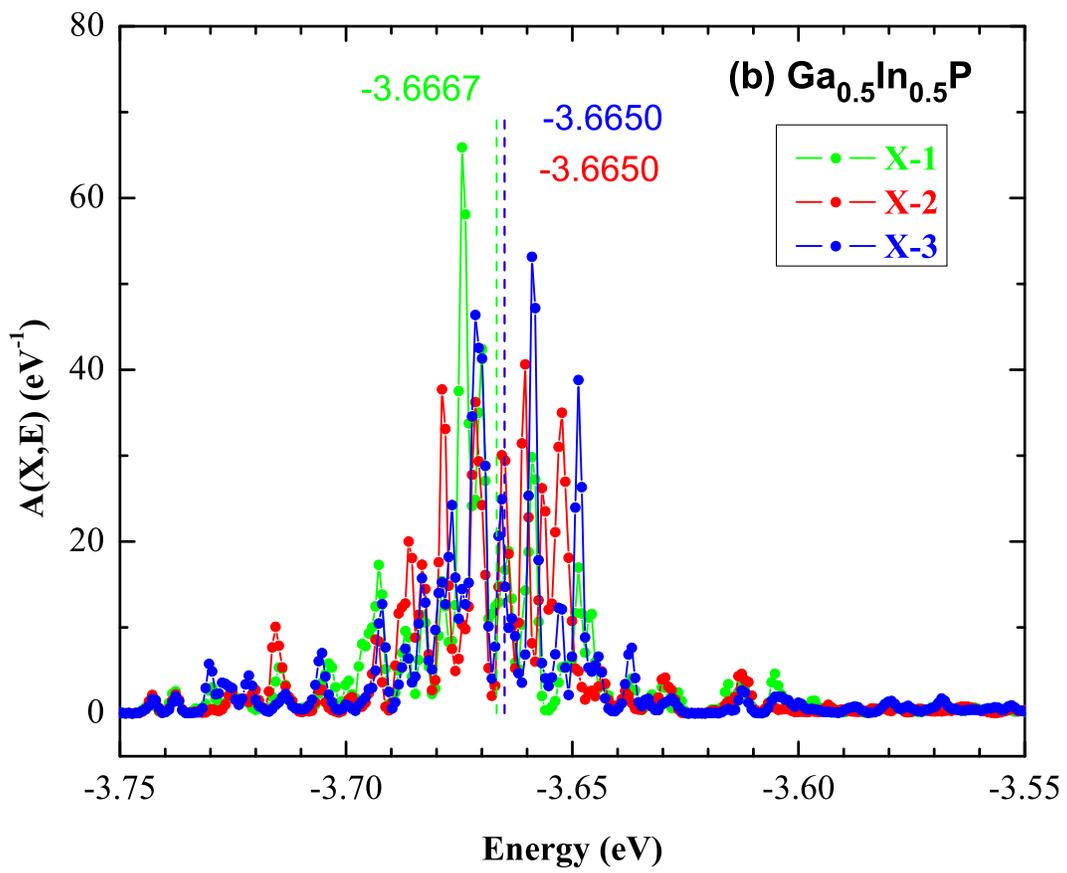


Figure 6b

BB11500

03MAR2011

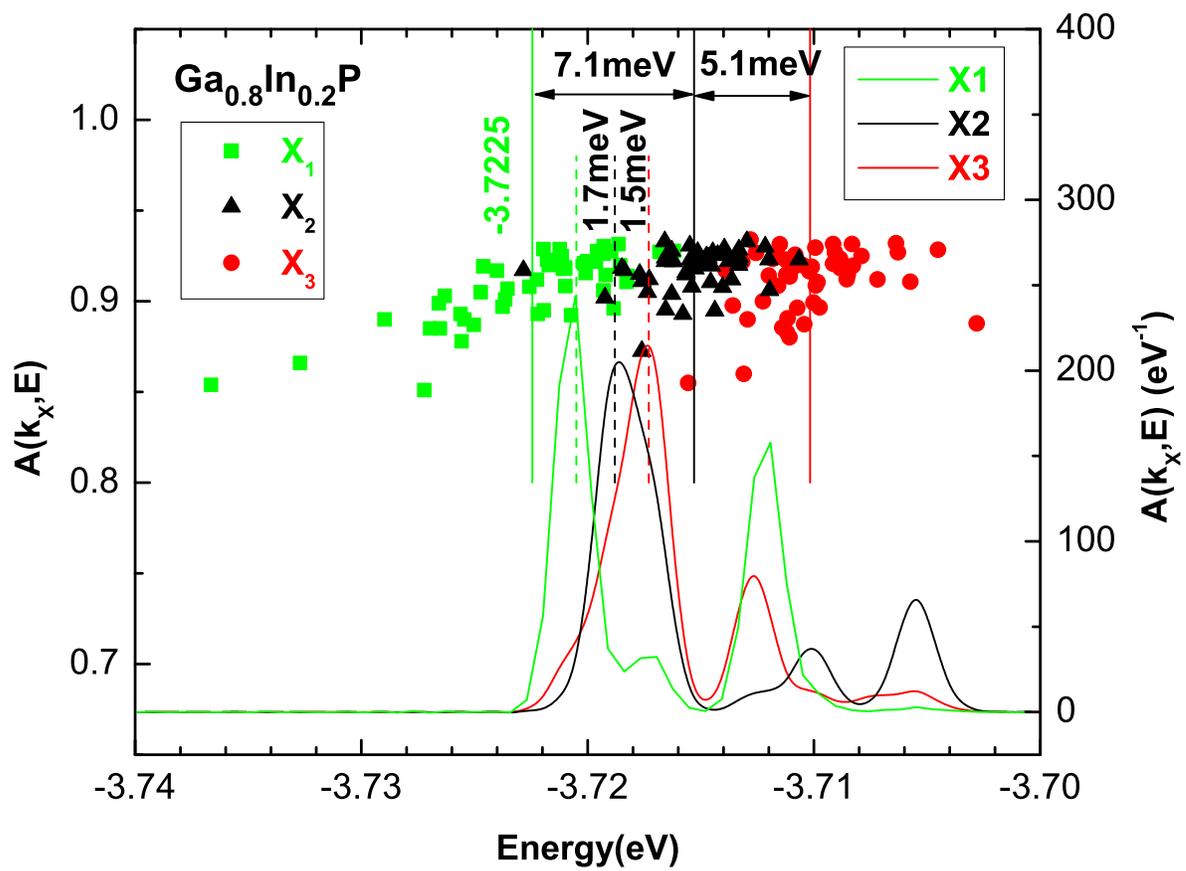


Figure 7

BB11500 03MAR2011