Independent and correlated composition behavior of material properties: Application to energy band gaps for the $Ga_\alpha In_{1-\alpha}P_\beta As_{1-\beta}$ and $Ga_\alpha In_{1-\alpha}P_\beta Sb_\gamma As_{1-\beta-\gamma}$ alloys

Kyurhee Shim
Department of Chemistry, Princeton University, Princeton, New Jersey 08544
and Department of Physics, Kyonggi University, Suwon 440-760, Korea

Herschel Rabitz
Department of Chemistry, Princeton University, Princeton, New Jersey 08544
(Received 15 December 1997; revised manuscript received 4 March 1998)

A correlated function expansion (CFE) is introduced (a) to identify the role of independent and correlated composition variations upon a desired material property, and (b) to provide an efficient means to compute the property throughout the composition space. As an example the contributions of independent and correlated composition behavior upon the principal energy band gaps for the alloys $Ga_\alpha In_{1-\alpha}P_\beta As_{1-\beta}$ and $Ga_\alpha In_{1-\alpha}P_\beta Sb_\gamma As_{1-\beta-\gamma}$ are calculated and analyzed by applying the CFE to the universal tight-binding (UTB) Hamiltonian model of the alloys. The convergence properties of the CFE over the entire composition variable space (α,β,γ) are examined upon including independent, pair-, and triple-correlated terms. By retaining only independent component contributions in the CFE it was possible to represent the UTB results to better than 90% accuracy for both the alloys $Ga_\alpha In_{1-\alpha}P_\beta As_{1-\beta}$ and $Ga_\alpha In_{1-\alpha}P_\beta Sb_\gamma As_{1-\beta-\gamma}$. Pair composition correlations contributed approximately 5–10 % to the band gaps in both alloys and for $Ga_\alpha In_{1-\alpha}P_\beta Sb_\gamma As_{1-\beta-\gamma}$ the triple correlations were at the level of ~3%. The CFE is a generic tool capable of simplifying efforts at finding desired alloy compositions for material properties. [S0163-1829(98)07028-3]

I. INTRODUCTION

Recently there has been increasing interest in the structural and electronic properties of the III-V semiconductor alloys and consequently in their electro-optical applications as high-efficiency light-emitting diodes and high-speed switching devices.[1–3] The independent and correlated behavior of the alloy components determines the electro-optical properties and give rise to nonlinear phenomena. However, no theoretical analysis has been reported on the effects of the independent and correlated component behavior in the III-V alloys with the number of components N satisfying $N \geqslant 2$. This situation is apparently due to the computational difficulties and complexity of dealing with disorder in the alloys, although marked differences with composition variation have been observed.[4–8] Extensive and systematic searches for desirable alloy compositions must be pursued to develop new materials. This effort will likely require new theoretical concepts and mathematical tools to provide the necessary physical insight and guidance to accelerate the laboratory efforts.

In this paper, a correlated function expansion (CFE) is introduced based on a high-dimensional model representation technique[9] to identify the independent and correlated composition behavior of multicomponent materials. The CFE is applied to investigate the contributions of independent and correlated composition behavior upon the Γ, L, and X energy band gaps for the alloys $Ga_\alpha In_{1-\alpha}P_\beta As_{1-\beta}$ and $Ga_\alpha In_{1-\alpha}P_\beta Sb_\gamma As_{1-\beta-\gamma}$ based on the universal tight-binding (UTB) model.[10] By utilizing a judiciously chosen *subset* of alloy compositions, the CFE can deduce the Γ, L, and X energy gaps for the *entire* composition variable space. The convergence properties of the CFE as a predictive tool are explored by examining the absolute error for the first- and second-order CFE expansion for 121 grid points (i.e., compositions) for the $Ga_\alpha In_{1-\alpha}P_\beta As_{1-\beta}$ alloys and 726 grid points for the $Ga_\alpha In_{1-\alpha}P_\beta Sb_\gamma As_{1-\beta-\gamma}$ alloys. Here, first- and second-order refer, respectively, to the independent and pair-correlated behavior amongst the composition variables. We found that the first-order CFE prediction is sufficient to represent the full composition space UTB theoretical values to over 90% accuracy. At this level of accuracy the first-order CFE predictions for both alloys correspond to computational savings of approximately factors of 6 and 40 for the respective alloys $Ga_\alpha In_{1-\alpha}P_\beta As_{1-\beta}$ and $Ga_\alpha In_{1-\alpha}P_\beta Sb_\gamma As_{1-\beta-\gamma}$. Small, but significant, pair-correlation composition behavior was evident. For the $Ga_\alpha In_{1-\alpha}P_\beta As_{1-\beta}$ alloy 100% accuracy occurs at second order with the pair correlations contributing approximately 5–10 % to the Γ band-gap value. In the $Ga_\alpha In_{1-\alpha}P_\beta Sb_\gamma As_{1-\beta-\gamma}$ system there were similar composition correlations as the CFE with terms up to second order predicted all the band gaps to better than 95% accuracy. As an example, in the latter alloy the independent, pair, and triple composition correlation contributions to the Γ band gap were, respectively, at the levels of 90%, 5.2%, and 4.8%. These CFE results have significant implications for simplifying alloy composition design efforts, as well as for analogous broader applications in materials science.[11]

II. CFE FOR MATERIALS DESIGN

The material property of interest (e.g., a band gap) is expressed as $\xi(x)$ where $x=\{x_1,x_2,\ldots,x_N\}$ is the collection of N component fractions. We seek a systematic and exact formulation for $\xi(x)$, which can identify the key role of each

component x_i with the aim of providing a basis to more efficiently determine useful new material compositions. Problems of this type have the apparent character of being nonpolynomial (NP) complete[12] to scale exponentially in computational effort $\sim S^N$, where S sample values are taken for each composition variable. The CFE can convert this task to only polynomic scaling in N and can also clearly identify the independent and correlated roles of the composition variables. Furthermore, the CFE technique does not employ regression analysis[13] and it permits arbitrary structure to exist in the composition space of the property $\xi(x)$.

In the CFE, the model output property for a multicomponent system $\xi(x)=\xi(x_1,x_2,\ldots,x_N)$ is expressed as a hierarchical correlated function expansion in terms of the input composition variables,

$$\begin{aligned}\xi(x)&=\xi_0+\sum_{i=1}^{N}\xi_i(x_i)+\sum_{1\leqslant i<j\leqslant N}\xi_{ij}(x_i,x_j)\\&\quad+\cdots+\xi_{123,\ldots,N}(x_1,x_2,\ldots,x_N)\\&\equiv\sum_{l=0}^{N}\eta_l(\hat{x}_l).\end{aligned}\tag{1}$$

Here, ξ_0 is a constant, $\xi_i(x_i)$ describes the independent role of the ith component, $\xi_{ij}(x_i,x_j)$ gives the correlated action of the variables x_i and x_j, etc. The CFE has a finite number of terms and is always an *exact* representation of $\xi(x)$ when sufficient terms are included (i.e., the last term picks up any final residual Nth-order correlations completing the description). The lth function $\eta_l(\hat{x}_l)$ represents a sum of all possible correlated functions at the lth order such that

$$\eta_l(\hat{x}_l)\equiv\sum_{1\leqslant i_1<i_2<i_l\leqslant N}\xi_{i_1i_2\ldots i_l}(x_{i_1},x_{i_2},\ldots,x_{i_l}),$$

where $\hat{x}_l\equiv\{x_{i_1}x_{i_2}\cdots x_{i_1}\}\subset x$. For example, $\eta_0=\xi_0$, $\eta_1(\hat{x}_1)=\Sigma_i\xi_i(x_i)$, $\eta_2(\hat{x}_2)=\Sigma_{i<j}\xi_{ij}(x_i,x_j)$, etc. Each member of the lth-order collection of correlated functions $\eta_l(\hat{x}_l)$ can be determined by a corresponding projection operation over the original function given by $\xi_{i_1i_2\cdots i_l}(\hat{x}_l)=\int P_{i_1i_2\cdots i_l}(\hat{x}_l;x)\xi(x)dx$, where the integral is over all N composition variables and $P_{i_1i_2\cdots i_l}(\hat{x}_l;x)$ is one member of the set of lth-order projection operators

$$\sum_{1\leqslant i_1<i_2<i_l\leqslant N}P_{i_1i_2\ldots i_l}(\hat{x}_l;x)\equiv P_l(\hat{x}_l;x),$$

which has the properties of othogonality and idempotency. The full set of such operators $P_{i_1i_2\cdots i_l}(\hat{x}_l;x)$ for $0\leqslant l\leqslant N$ provides a resolution of the identity operator, and the output property function $\xi_{i_1i_2\cdots i_l}(\hat{x}_l)$ is invariant under the associated projection operator $P_l(\hat{x}_l,x)$. Thus, each correlated function $\xi_{i_1i_2\cdots i_l}(\hat{x}_l)$ provides independent and unique physical information on the relationships among the input composition variables for their action upon the output material property. Furthermore, no restrictions are placed on the functional form of the correlation terms, thereby admitting highly nonlinear material properties with respect to the constituent composition. Thus, for example, a truncation of the CFE even to first order [i.e., $\xi(x)\sim\eta_0+\eta_1(\hat{x}_1)$] is generally not a linear combination of fractional composition contributions due to the nonlinear nature of the functions $\xi_i(x_i)$. The concept of material composition correlations introduced here is not statistically based, but rather its origin lies in the inherent behavior of the material property composition space.

Various choices for the projection operators are possible,[9] but the following one is especially useful,

$$\begin{aligned}P_{i_1i_2\cdots i_l}(\hat{x}_l;x')&=\prod_{s=1}^{l}\left[\delta(x'_{i_s}-x_{i_s})-\delta(x'_{i_s}-\overline{x_{i_s}})\right]\\&\quad\times\prod_{k=l+1}^{N}\delta(x'_{1_k}-\overline{x_{i_k}}),\quad l\geqslant 1,\end{aligned}\tag{2}$$

with the additional zeroth-order projection operator

$$P_0=\prod_{i=1}^{N}\delta(x_i-\overline{x_i}).$$

These projection operators will uniquely specify all the correlated functions order by order, where $\overline{x}=(\overline{x_1},\overline{x_2},\ldots,\overline{x_N})$ is a chosen fixed composition point in the variable space, called a cut center. It is implicitly understood that P_l denotes one member of a set of projection operators at the lth order as specified by the choice of the components constituting $\hat{x}_l$. The cut center $\overline{x}$ would typically be a chosen reference material around which composition explorations (i.e., perhaps large, including the entire space) would be made. When the CFE is taken out to convergence, then the choice of $\overline{x}$ is irrelevant. Operating with the projection operator defined in Eq. (2) upon the original material function results in the following CFE terms:

$$\begin{aligned}\xi_0&=\xi(\overline{x}),\\\xi_i(x_i)&=\xi(x_i;\overline{x^i})-\xi_0,\\\xi_{ij}(x_ix_j)&=\xi(x_i,x_j;\overline{x^{ij}})-\xi_i(x_i)-\xi_j(x_j)-\xi_0,\end{aligned}\tag{3}$$

etc. Here, ξ_0 is the model output material property for the all of the composition variables fixed at the reference value $\overline{x}$. The term $\xi(x_i,\overline{x^i})$ is the model output material property as a function of the single component x_i (i.e., a cut along x_i through the cut center in the composition space), while the other variables, $x_j\equiv\overline{x_j}$, $j\neq i$, are fixed at the cut center. Therefore, the behavior of $\xi_i(x_i)$ determines the property response upon the independent action of the material component variable x_i with respect to the reference system. In the same manner, $\xi(x_i,x_j;\overline{x^{ij}})$ indicates the model output material property for all the composition variables, $x_k\equiv\overline{x_k}$, $k\neq i,j$, fixed at the cut center except for x_i and x_j. Thus, $\xi_{ij}(x_i,x_j)$ describes the pair-correlated behavior between the variables x_i and x_j with respect to the reference material identified at the cut center. A similar interpretation would apply to additional higher-order CFE terms. In this manner, the CFE operational procedure determines all the correlated composition behavior with each term having a clear physical meaning on the right-hand side of Eq. (1) tailored to any complex material property $\xi(x)=\xi(x_1,x_2,\ldots,x_N)$.

In most physical (i.e., material) systems, one anticipates that higher-order correlations are unlikely to be significant.

This implies that the CFE is expected to converge at low order and thereby to save in computational efforts at exploring the composition space. There is considerable evidence supporting rapid convergence from statistics for many other problems where only low-order (e.g., up to second order, typically) multivariable variances are adequate, and recently the fast convergence of the CFE has been observed in atmospheric research.[14] This paper also shows that the CFE taken to just first order can provide good accuracy to describe the composition variation of the energy band gaps of the alloys $Ga_\alpha In_{1-\alpha}P_\beta As_{1-\beta}$ (Ref. 10) and $Ga_\alpha In_{1-\alpha}P_\beta Sb_\gamma As_{1-\beta-\gamma}$ (Ref. 15) based on the UTB model.

Operationally the material property $\xi(\tilde{x})$ at an arbitrary point $\tilde{x}$ in the component space would be evaluated by projecting $\tilde{x}$ onto the various contributing cuts (i.e., lines, surfaces, volumes, etc.) and performing the associated low-dimensional interpolations on the cuts. The latter interpolations are extremely fast to perform. Although sample cuts through a high-dimensional composition space are natural to consider, without the formal guiding structure of the CFE it would not be evident how to employ the cuts for a global determination of $\xi(\tilde{x})$.

The computational savings afforded by the CFE can easily be estimated. If the CFE converges at mth order with acceptable accuracy and considering a sampling of S composition values in each variable, then the total number of operations needed to calculate the CFE is (considering the number of correlated functions at each order)

$$\sum_{k=0}^{m} \frac{N!}{(N-k)!k!}(S-1)^k. \tag{4}$$

It was assumed that the cut center is at one of the sampled composition values. When taking this result to the extreme limit of the CFE requiring all terms up to Nth order, we find that the total number of sampling operations is S^N $[=\Sigma_{k=0}^{N}[N!/(N-k)!k!](S-1)^k]$. This limit is exactly the exponentially difficult NP complete scaling that one would normally expect without consideration of the unlikelihood of finding high-order correlations. Therefore, the CFE can do no worse than S^N scaling, and on physical grounds we expect a dramatic savings with evidence suggesting that $m \ll N$. For example, consider an $N=8$ variable physical system sampled with $S=10$ points per variable and assume that the CFE yields acceptable convergence (e.g., to a few percent error as indicated in Sec. III) at second order, then the ratio of the computational effort of CFE to that of conventional NP scaling is

$$\left\{\sum_{k=0}^{2} \frac{N!}{(N-k)!k!}(S-1)^k\right\} \Big/ \{S^N\} \sim 10^{-5}.$$

Thus, the CFE operation can render the original exponentially difficult computations down to a problem of only polynomial complexity. The substantiation of this dramatic computational savings could have a profound impact on material design efforts at finding new materials in the laboratory. Another measure of savings is the cost of calculating the material property at a new point $\tilde{x}$ by the CFE relative to the effort by the electronic structural code. The savings here can be equally as dramatic, since the CFE evaluated at low order with interpolation over its stored grid is a very rapidly performed calculation. Furthermore, the extracted individual CFE terms provide clear physical insight and guidance on the composition correlations contributing to the desired physical and/or chemical properties. The illustrations in the following section are modest in complexity, but they do support this key low-order CFE convergence behavior already demonstrated in other physical systems of very high dimensions.[9]

III. APPLICATION TO THE $Ga_\alpha In_{1-\alpha}P_\beta As_{1-\beta}$ AND $Ga_\alpha In_{1-\alpha}P_\beta Sb_\gamma As_{1-\beta-\gamma}$ ALLOYS

The independent and correlated composition variable dependence of the principal energy band gaps E in the III-V alloys (here, $Ga_\alpha In_{1-\alpha}P_\beta As_{1-\beta}$ and $Ga_\alpha In_{1-\alpha}P_\beta Sb_\gamma As_{1-\beta-\gamma}$) can be identified by applying the CFE (in Sec. II) to the UTB alloy Hamiltonian.[10,15] In UTB, the alloy Hamiltonian is represented by an effective matrix consisting of both ordered and disordered composition-dependent components induced by the atomic orbital interaction.[10] The composition-dependent energy band gaps are obtained by diagonalizing the alloy Hamiltonian matrix. The composition variables (α,β,γ) are embedded in the Hamiltonian matrix elements as input parameters and the composition dependence of energy band structure and band gaps are the output material properties obtained through the diagonalization of the alloy Hamiltonian. The CFE application to the UTB theory will determine the band-gap $E(\alpha,\beta,\gamma)$ dependence upon the material component variables (i.e., α,β,γ) through their independent $(E_\alpha,E_\beta,E_\gamma,\ldots)$, pair- $(E_{\alpha\beta},E_{\beta\gamma},E_{\alpha\gamma},\ldots)$, and triple-$(E_{\alpha\beta\gamma})$ correlated actions with respect to the reference material at the cut center $(\bar{\alpha},\bar{\beta},\bar{\gamma})$.

The energy band gaps of the reference alloys $Ga_{\bar{\alpha}}In_{1-\bar{\alpha}}P_{\bar{\beta}}As_{1-\bar{\beta}}$ and $Ga_{\bar{\alpha}}In_{1-\bar{\alpha}}P_{\bar{\beta}}Sb_{\bar{\gamma}}As_{1-\bar{\beta}-\bar{\gamma}}$ are obtained by fixing all the composition variables at the reference value $(\bar{\alpha},\bar{\beta},\bar{\gamma})$ such that $E_0 \equiv \langle \hat{H}(\bar{\alpha},\bar{\beta},\bar{\gamma})\rangle$, where $\langle \hat{H}\rangle$ denotes the diagonalization of the Hamiltonian to obtain the band gap. The independent composition behavior $(E_\alpha,E_\beta,E_\gamma)$ response to each composition variable with respect to the reference material is determined from Eq. (3) as follows: $E_\alpha \equiv \langle \hat{H}(\alpha,\bar{\beta},\bar{\gamma})\rangle - E_0$, etc. In the same manner, the pair- and triple-correlated composition behaviors $(E_{\alpha\beta},\ldots,E_{\alpha\beta\gamma})$ can be obtained from Eq. (3) such that $E_{\alpha\beta} \equiv \langle \hat{H}(\alpha,\beta,\bar{\gamma})\rangle - E_\alpha - E_\beta - E_0$, and $E_{\alpha\beta\gamma} \equiv \langle \hat{H}(\alpha,\beta,\gamma)\rangle - E_{\alpha\beta} - E_{\alpha\gamma} - E_{\beta\gamma} - E_\alpha - E_\beta - E_\gamma - E_0$, etc. The total composition dependence of an energy band gap is expressed by the CFE in Eq. (3) as

$$E(\alpha,\beta,\gamma) = E_0 + \sum_{i=\alpha,\beta,\gamma} E_i + \sum_{i<j=\alpha,\beta,\gamma} E_{ij} + E_{\alpha\beta\gamma}.$$

If the CFE operation is taken up to Nth order (cf. $N=2$ or 3 for the two alloys treated here), the energy band gap exactly equals that of the UTB model. There is no $E_{\alpha\beta\gamma}$ term for the $N=2$ alloy and the expansion is exact at second order in that case.

In order to examine the convergence of the CFE up to mth order (E^m), we define the CFE partial sums

$$E^0 \equiv E_0, \quad E^1 \equiv E_0 + \sum_{i=\alpha,\beta,\gamma} E_i,$$

TABLE I. Independent and correlated composition band-gap behavior of $Ga_\alpha In_{1-\alpha}P_\beta Sb_\gamma As_{1-\beta-\gamma}$. The calculated independent ($E_\alpha, E_\beta, E_\gamma$), pair- ($E_{\alpha\beta}, E_{\beta\gamma}, E_{\alpha\gamma}$), and triple- ($E_{\alpha\beta\gamma}$) correlated values and the zeroth- (E^0), first- (E^1), and second-order (E^2) CFE results along with the UTB value (E_{UTB}) for the principal energy band gaps at three different compositions (α,β,γ) for the $Ga_\alpha In_{1-\alpha}P_\beta Sb_\gamma As_{1-\beta-\gamma}$ alloy [$a \equiv (0,0,0)$:InAs; $b \equiv (0.2,\ 0.2,\ 0.2)$:$Ga_{0.2}In_{0.8}P_{0.2}Sb_{0.2}As_{0.6}$; and $c \equiv (0.4,\ 0.4,\ 0.4)$:$Ga_{0.4}In_{0.6}P_{0.4}Sb_{0.4}As_{0.2}$]. All units are in eV.

$\vec{r}$:Gap	E_α	E_β	E_γ	$E_{\alpha\beta}$	$E_{\beta\gamma}$	$E_{\alpha\gamma}$	$E_{\alpha\beta\gamma}$	E^0	E^1	E^2	E_{UTB}
a:Γ	−0.467	−0.332	0.161	0.074	0.077	−0.047	0.075	0.830	0.192	0.296	0.370
a:L	−0.210	−0.208	0.212	0.054	0.039	0.035	0.028	1.553	1.347	1.475	1.502
a:X	0.055	0.079	0.233	−0.072	−0.004	0.042	0.063	1.884	2.251	2.217	2.280
b:Γ	−0.310	−0.122	0.037	0.009	0.009	−0.017	0.005	0.830	0.435	0.436	0.441
b:L	−0.138	−0.075	0.060	0.011	0.005	0.001	0.002	1.553	1.400	1.417	1.420
b:X	0.024	0.026	0.067	−0.012	∼0	0.003	0.004	1.884	2.002	1.992	1.996
c:Γ	−0.114	0.134	−0.021	∼0	0.01	0.009	0.002	0.830	0.829	0.847	0.849
c;L	−0.050	0.078	−0.049	−0.004	0.004	0.002	0.002	1.553	1.533	1.535	1.537
c:X	0.005	−0.025	−0.056	0.003	∼0	∼0	0.002	1.884	1.809	1.813	1.814

$$E^2 \equiv E_0 + \sum_{i=\alpha,\beta,\gamma} E_i + \sum_{i<j=\alpha,\beta,\gamma} E_{ij},$$

etc. The absolute fractional error ϵ_i^m at an arbitrary point $i \equiv (\alpha_i, \beta_i, \gamma_i)$ for the mth-order CFE in relation to the true model output value E_i is defined as $\epsilon_i^m = |E^m - E(\alpha_i,\beta_i,\gamma_i)| / E(\alpha_i,\beta_i,\gamma_i)$. The average error ($\overline{\epsilon_m}$) over the whole composition variable space is

$$\overline{\epsilon^m} = (1/n) \sum_{i}^{\text{all space}} \epsilon_i^m,$$

where n is the total number of composition grid points in the variable space: in the present cases $n=121$ for $Ga_\alpha In_{1-\alpha}P_\beta As_{1-\beta}$ and $n=726$ for $Ga_\alpha In_{1-\alpha}P_\beta Sb_\gamma As_{1-\beta-\gamma}$. The average error provides an overall measure of the convergence of the CFE.

IV. RESULTS AND DISCUSSION

The energy band structure and band gaps at arbitrary component compositions for $Ga_\alpha In_{1-\alpha}P_\beta As_{1-\beta}$ ($N=2;\alpha,\beta$) and $Ga_\alpha In_{1-\alpha}P_\beta Sb_\gamma As_{1-\beta-\gamma}$ ($N=3;\alpha,\beta,\gamma$) have been calculated by the UTB method. Applying the CFE to the UTB model revealed the independent and correlated behavior of the three principal energy band gaps [$E(\Gamma)$, $E(L)$, and $E(X)$] for both systems. The TB parameters and bond lengths for the six constituent binaries (InP, InAs, InSb, GaP, GaAs, and GaSb) were taken from Refs. 10, 16, and 17. The cut center was chosen to be roughly in the middle of the composition variable space, so that $(\overline{\alpha},\overline{\beta}) = (0.5, 0.5)$ for $Ga_\alpha In_{1-\alpha}P_\beta As_{1-\beta}$ and $(\overline{\alpha},\overline{\beta},\overline{\gamma}) = (0.5, 0.3, 0.3)$ for $Ga_\alpha In_{1-\alpha}P_\beta Sb_\gamma As_{1-\beta-\gamma}$. Although picking the center of the space is reasonable, little sensitivity was found with respect to the choice (i.e., the CFE is invariant to the choice at full convergence). Each order of CFE expansion was tested over the entire range of composition space as discussed above.

In Table I, the independent, pair-, and triple-correlated composition behaviors of the $E(\Gamma)$, $E(L)$, and $E(X)$ band gaps predicted by the CFE are shown and compared with the true UTB output model value for three sampled compositions of the $Ga_\alpha In_{1-\alpha}P_\beta Sb_\gamma As_{1-\beta-\gamma}$ alloy along the [1,1,1] variable direction from the origin (0,0,0). The absolute magnitude of the independent (E_i), pair- (E_{ij}), and triple- (E_{ijk}) correlated contributions to the energy band gaps generally decreases with order such that $|E_i| \gg |E_{ij}| > \sim |E_{ijk}|$, while their sign (+ or −) indicates that the contribution accordingly acts to increase or decrease the band gaps. Also, the values predicted by the CFE are excellent even at first order especially as the composition variable gets close to the cut center. The CFE converges rapidly to accurate values in all cases with the exception of the $E(\Gamma)$ gap at the (0,0,0) composition, but even here the accuracy is better than 80% at second order. In general we find that although there are some isolated regions of moderate accuracy, the average error over the entire composition space for all the band gaps is excellent (cf. Table II discussed later).

Figures 1 and 2 display examples of the independent and pair-correlated composition contributions for the $Ga_\alpha In_{1-\alpha}P_\beta As_{1-\beta}$ and $Ga_\alpha In_{1-\alpha}P_\beta Sb_\gamma As_{1-\beta-\gamma}$ alloys. Figure 1 shows that the components acting independently (i.e., the individual action of Ga, P, and Sb in the alloys) play different roles for the distinct band gaps: in some cases the energy gaps increase while in other cases they decrease, and generally nonlinearly, with component fraction. The independent behavior depends on the type of band gap and the specific composition region. A noticeable feature of the independent behaviors is that they are not linear. Linear variable behavior has been often assumed in many approximations such as the virtual crystal approximation (VCA). However, extensive experimental reports on the band-gap variation for the III-V ternary alloys $Al_xGa_{1-x}As$, $Ga_xIn_{1-x}P$, $In_xGa_{1-x}As$, $In_xAs_{1-x}P$, etc. exhibit nonlinear bowing behavior.[6–8] The $E(\Gamma)$ band gaps of the $In_{1-\alpha}Ga_\alpha As$ and $InAs_{1-\beta}P_\beta$ ternary alloys predicted by the first-order CFE are plotted in Fig. 3, respectively, as the $\beta=0$ or $\alpha=0$ limits for the $Ga_\alpha In_{1-\alpha}P_\beta As_{1-\beta}$ alloy. Comparing the calculations with the experimental data in the figure shows that there is excellent agreement for the $In_{1-\alpha}Ga_\alpha As$ alloy and good agreement for the $InAs_{1-\beta}P_\beta$ alloy. These results support the quality of the UTB calculations and show that the first-order CFE is sufficient to represent the band gaps to good accuracy. In the

TABLE II. Overall percent accuracy $(1-\overline{\epsilon^m})\times 100\%$ for the band gaps from the zeroth-, first-, and second-order CFE predictions over the entire composition variable space.

Alloy system	Gap	$(1-\overline{\epsilon^0})\times 100(\%)$	$(1-\overline{\epsilon^1})\times 100(\%)$	$(1-\overline{\epsilon^2})\times 100(\%)$	References (eV)[a]
$Ga_\alpha In_{1-\alpha}P_\beta As_{1-\beta}$	Γ	53.6	95.4	100	$E^0(\Gamma)=1.218$
					$0.37<E(\Gamma)<2.88$
	L	87.6	98.2	100	$E^0(L)=1.905$
					$1.60<E(L)<2.72$
	X	96.5	98.5	100	$E^0(X)=2.071$
					$1.97<E(X)<2.18$
$Ga_\alpha In_{1-\alpha}P_\beta Sb_\gamma As_{1-\beta-\gamma}$	Γ	27.8	90.0	94.8	$E^0(\Gamma)=0.830$
					$0.25<E(\Gamma)<2.88$
	L	82.9	97.3	97.9	$E^0(L)=1.553$
					$0.97<E(L)<2.72$
	X	93.8	98.3	97.7	$E^0(X)=1.884$
					$1.69<E(X)<2.28$

[a]The reference value and range of the band gaps are indicated.

UTB model, the independent composition behavior of the energy band gaps can be analyzed with respect to the reference material, in which the nonlinearity of the band gap is induced by the atomic orbital electronic interaction of the disordered units through bond alternation. The addition of Sb on the anion site in the alloy leaves the independent behavior of Ga and P similar to that found in the $N=2$ case. The band-gap response to the total action of all the independent variables could be simple or complex depending on the composition region and band-gap type $[E(\Gamma), E(L),$ and $E(X)]$.

Examples of the pair correlated composition behaviors are shown in Fig. 2. These pair terms are associated with second-order atomic orbital interactions between the components already disordered through bond alternation. They will readjust the conduction-band minimum level as a function of composition to some degree and especially in terms of the band-gap bowing. The correlated function surfaces have considerable structure according to the combination of variables, variable range, and band-gap type. Comparing the correlated behaviors for $N=3$ with that for $N=2$, we notice that the absolute magnitude of the correlated behavior decreases with the addition of the Sb component. This result implies that generally the independent component behavior in the

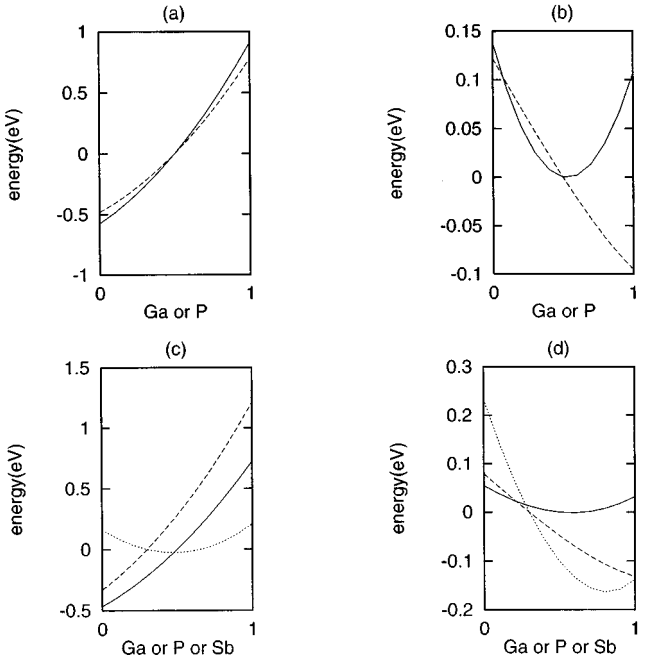

FIG. 1. Examples of independent composition behavior; (a) $E_i(\Gamma)$ for the $Ga_\alpha In_{1-\alpha}P_\beta As_{1-\beta}$ alloy, (b) $E_i(X)$ for the $Ga_\alpha In_{1-\alpha}P_\beta As_{1-\beta}$ alloy, (c) $E_i(\Gamma)$ for the $Ga_\alpha In_{1-\alpha}P_\beta Sb_\gamma As_{1-\beta-\gamma}$ alloy, and (d) $E_i(X)$ for the $Ga_\alpha In_{1-\alpha}P_\beta Sb_\gamma As_{1-\beta-\gamma}$ alloy. Here, $i=$Ga is the solid line, $i=$P is the dashed line, and $i=$Sb is the dotted line.

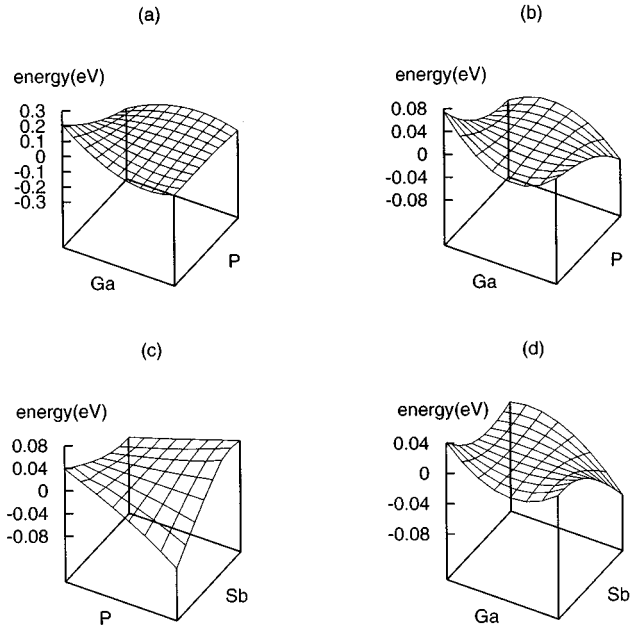

FIG. 2. Examples of pair-correlated composition behavior; (a) $E_{\alpha\beta}(\Gamma)$ for the $Ga_\alpha In_{1-\alpha}P_\beta As_{1-\beta}$ alloy, (b) $E_{\alpha\beta}(\Gamma)$ for the $Ga_\alpha In_{1-\alpha}P_\beta Sb_\gamma As_{1-\beta-\gamma}$ alloy, (c) $E_{\beta\gamma}(L)$ for the $Ga_\alpha In_{1-\alpha}P_\beta Sb_\gamma As_{1-\beta-\gamma}$ alloy, and (d) $E_{\alpha\gamma}(X)$ for the $Ga_\alpha In_{1-\alpha}P_\beta Sb_\gamma As_{1-\beta-\gamma}$ alloy.

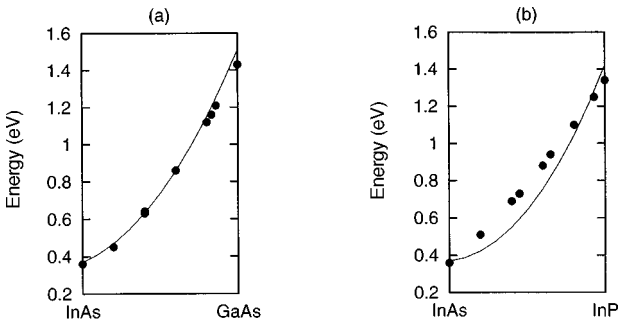

FIG. 3. The composition $E(\Gamma)$ gap variation by the first-order CFE (solid line) E^1 for the ternary alloys (a) $Ga_\alpha In_{1-\alpha}As$, (b) $InP_\beta As_{1-\beta}$ compared with the experimental data (●) (Ref. 6).

$Ga_\alpha In_{1-\alpha}P_\beta Sb_\gamma As_{1-\beta-\gamma}$ alloy ($N=3$) is more dominant than in the $Ga_\alpha In_{1-\alpha}P_\beta As_{1-\beta}$ alloy ($N=2$). This observation holds, despite the fact that there are more second-order terms for $N=3$ than $N=2$. The absolute magnitude of the correlated behavior increases when moving further away from the cut reference frame and similar effects are evident in the independent behavior plots. Note that along any of the cuts the associated second-order surfaces are zero. This property indicates that the CFE should be very accurate around the cuts and especially the cut center. The exact value of a pair-correlated contribution depends on the particular variables, their region, and band-gap type. The independent contributions of the first-order terms can either reinforce or detract from what is found in the second-order terms. By examining the collective behaviors of all the terms one can deduce the various combinations of components that give desirable material properties (here, the band gap). This type of composition identification is a very important issue for these semiconductors as well as for other analogous materials applications.[11]

The average percent accuracy, $(1-\overline{\epsilon^m})\times 100\%$ defined in Sec. III is shown in Table II. The results for the average accuracy versus CFE order $m=0$, 1, and 2 for the $E(\Gamma)$, $E(L)$, and $E(X)$ gaps give a simple global measure of the CFE quality. Excellent accuracy ($\geqslant 90\%$) occurs in the vicinity of cut center even for the first-order CFE, but larger errors (more than 10%) occur around the edge of the variable space even for the second-order CFE. For the $Ga_\alpha In_{1-\alpha}P_\beta Sb_\gamma As_{1-\beta-\gamma}$ alloy ($N=3$), the $E(\Gamma)$ gap shows the most dramatic improvement order by order when we consider the wide range of the $E(\Gamma)$ gap variation $[0.25 < E(\Gamma) < 2.88]$ over the entire composition space. The zeroth-, first-, and second-order CFE for the $E(\Gamma)$ gap predicted the UTB values with an accuracy of ~28%, ~90%, and ~95%, respectively. Thus, the band gaps can be rapidly predicted over a wide scale by a low-order CFE to very good accuracy. This is a very important aspect of the CFE when we consider that the computational savings of approximately factors of 6 for $N=2$ and of 40 for $N=3$ result from using the first-order CFE. Furthermore, for alloys with large numbers ($N\gg 1$) of components the computational savings can be dramatic as indicated in Sec. II. For the $E(L)$ and $E(X)$ band gaps, the CFE reproduced the UTB values over the full space to 97–98 % accuracy even at the first-order level. For all the band gaps with a first-order CFE the gap standard deviations (not shown) are quite small over the entire composition range. Considering all the band gaps over the full composition space, it was found that the pair and triple (for $N=3$) component correlations contribute 5–10 % and ~3%, respectively. The accuracy of the first-order CFE depends on the range of variation of the output material property; i.e., a smaller range yields better accuracy at any particular CFE order. In one isolated location ($x\sim 1$ and $y\sim 1$) we found that the $E(X)$ band gap for the second-order CFE had lower accuracy (~80%) than the first-order CFE. This results in the average error at second order in Table II apparently showing no improvement over first order; however, in most of the other regions the accuracy of the second-order CFE is improved over the first-order CFE.

In order to examine the role of the cut center upon the average error $\overline{\epsilon}$, we performed another CFE for the $N=2$ case by taking the cut center off to one side at (0,0). The average accuracy with the first-order CFE was approximately 90% for $E(\Gamma)$, 95% for $E(L)$, and 97% for $E(X)$ band gaps. These results are only slightly less accurate than those found for the more natural cut center at (0.5, 0.5) throughout the full variable space. Thus, the choice of cut center is not a critical issue, but as expected better accuracy was found in the vicinity of the cut center for either case [e.g., near (0,0) where the composition is InAs]. Considering both of these cut centers, the independent and pair composition correlations contribute approximately 90–95 % and 5–10 %, respectively, to the $E(\Gamma)$ band gap in the $Ga_\alpha In_{1-\alpha}P_\beta As_{1-\beta}$ alloy ($N=2$). In practice a particular cut center would likely be chosen around a known interesting alloy composition for the desired material property.

V. CONCLUSION

The CFE was presented to identify the role of independent and correlated composition variations upon a desired material property. It was successfully applied to compute the energy band gaps for the $Ga_\alpha In_{1-\alpha}P_\beta As_{1-\beta}$ and $Ga_\alpha In_{1-\alpha}P_\beta Sb_\gamma As_{1-\beta-\gamma}$ alloys. The contributions of independent and correlated composition behavior to the principal energy band gaps $[E(\Gamma),E(L),E(X)]$ for these two alloys are obtained by applying the CFE to the UTB Hamiltonian model. From these investigations, we found the following. (1) The independent component behavior is not linear. (2) The band gaps in the region around the cut center were predicted well by the first-order CFE to over 95% accuracy. (3) The average accuracy was ~90% over the entire composition space for both alloys at just the first-order CFE. Thus, the first-order CFE is generally sufficient to represent the material band-gap properties. (4) The pair-correlated composition behavior for the $Ga_\alpha In_{1-\alpha}P_\beta As_{1-\beta}$ alloy ($N=2$) is reduced by adding the Sb component in the anion site [i.e., for the $Ga_\alpha In_{1-\alpha}P_\beta Sb_\gamma As_{1-\beta-\gamma}$ alloy ($N=3$)]. (5) The contributions of the independent, pair, and triple composition correlations were approximately at the levels of ~90%, 5–10 %, and ~3% in the $Ga_\alpha In_{1-\alpha}P_\beta Sb_\gamma As_{1-\beta-\gamma}$ alloy over the entire composition space. (6) All the independent and correlated behaviors have some variability according to the combination of components and band-gap type, but some similar behavior could be identified. As argued in Sec. II, the CFE

can greatly accelerate an exploration of the full composition space of the material. A related issue is the cost of evaluating the CFE at a new composition value (i.e., through interpolation among the stored CFE terms) vs that of calculating that composition by performing another electronic structure calculation. In the case of the UTB model and the second-order CFE, the savings by the CFE per evaluation was several orders of magnitude. Such savings open up the possibilities of systematic materials searches that were previously very difficult to perform.

Although the focus in this paper was on band gaps in semiconductors, the same CFE tools could be applied to other material properties. In the context of combinatorial materials synthesis[11] one could view the electronic wave functions evaluated and stored over the composition cuts as forming a library for various applications (naturally, the reliability of the CFE materials property predictions will depend on the quality of the wave functions making up the library). Importantly this library would have selected members designed by CFE concepts rather than a random sampling of the full space. Furthermore analogous tools can be applied directly in the laboratory to aid in the systematic identification of new materials.[18]

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation. One of the authors (H.R.) acknowledges many useful discussions on HDMR with Dr. J. Shorter.

[1] F. Devaux, S. Chelles, A. Ougazzaden, A. Mircea, and J. C. Harmand, Semicond. Sci. Technol. **10**, 887 (1995).
[2] M. Feng, C. L. Lau, V. Eu, and C. Ito, Appl. Phys. Lett. **51**, 1233 (1990).
[3] Maruice Quillec, *Materials for Optoelectronics* (Kluwer Academic, Boston, 1996).
[4] R. Benzaquen, S. Charbonneau, N. Sawadsky, A. P. Roth, R. Leonelli, L. Hobbs, and G. Knight, J. Appl. Phys. **75**, 2633 (1994).
[5] V. B. Gera, R. Gupta, and K. T. Jain, Phys. Rev. B **36**, 9657 (1987).
[6] Z. Q. Li and W. Potz, Phys. Rev. B **46**, 2109 (1992); A. P. Roth and E. Fortin, Can. J. Phys. **56**, 1468 (1978).
[7] C. Bosio, J. L. Staehli, M. Guzzi, G. Burri, and R. A. Logan, Phys. Rev. B **38**, 3263 (1988).
[8] O. Madelung, M. Schulz, and H. Weiss, *Numerical Data and Functional Relationships in Science and Technology: Semiconductors* (Springer-Verlag, Berlin, 1982), Vol. 17.
[9] O. F. Alis, H. Rabitz, K. Shim, and J. Shorter (unpublished).
[10] K. Shim and H. Rabitz, Phys. Rev. B **57**, 12 874 (1998).
[11] R. F. Service, Science **277**, 474 (1997).
[12] B. Barker, J. Assoc. Comput. Machin. **41**, 153 (1994).
[13] A. I. Khuri and J. A. Cornell, *Response Surfaces: Designs and Analyses* (Dekker, New York, 1996).
[14] H. Rabitz, P. Ip, and J. Shorter, J. Geophys. Res. (unpublished).
[15] K. Shim and H. Rabitz (unpublished).
[16] D. N. Talwar and C. S. Ting, Phys. Rev. B **25**, 2660 (1982).
[17] K. Shim, J. W. Bae, S. W. Jeong, H. J. Moh, and D. N. Talwar, Solid State Commun. **98**, 825 (1996).
[18] H. Rabitz and K. Shim (unpublished).