Ga$_{1-x}$Al$_x$N system, Madelung, and strain energies:
A study on the quality of cluster expansions

L. G. Ferreira
Instituto de Física Gleb Wataghin, Universidade Estadual de Campinas, Caixa Postal 6165, 13083-970 Campinas, SP, Brazil

M. Marques
Departamento de Microondas e Optoeletrônica, Divisão de Engenharia Eletrônica, Instituto Tecnológico de Aeronáutica, Comando-Geral de Tecnologia Aeroespacial, 12228-900 São José dos Campos, SP, Brazil

L. K. Teles
Departamento de Física, Instituto Tecnológico de Aeronáutica, Comando-Geral de Tecnologia Aeroespacial, 12228-900 São José dos Campos, SP, Brazil

(Received 6 February 2006; revised manuscript received 3 July 2006; published 18 August 2006; corrected 24 August 2006)

First-principles calculations within the local-density formalism were used to study the accuracy of cluster expansion techniques to predict the energy band gaps and enthalpy of the pseudobinary system Ga$_{1-x}$Al$_x$N, a technologically important alloy. The chosen pseudobinary system has the advantage of having small lattice mismatches, which minimizes the enthalpies of formation, and of being a semiconducting system with a direct band gap for any concentration $x$. Many different cluster expansion techniques were tested, some presenting clear advantages. The many cluster expansions were also compared against models of Madelung and strain energy, both long-range interactions. Though cluster expansions fail completely for the long-range Madelung interaction model, they behave remarkably well in the not so long-range strain model. The qualitative results for the strain model are similar to the results for the enthalpy and gap of the alloy system, thus giving us an assurance of our conclusions. Using only short-range interactions, all cluster expansions are clearly inadequate for the long-period orderings.

DOI: 10.1103/PhysRevB.74.075324 PACS number(s): 71.15.Nc, 71.20.Nr, 71.22.+i, 71.55.Eq

I. INTRODUCTION

Cluster expansions (CE’s) have a long history in the Ising model for alloy physics.1–4 Much is usually expected from them,5,6 sometimes without a secure theoretical justification. Presently we find batteries of programs and procedures aiming at applying and using CE’s in all sorts of alloy studies.7 New proposals are always arriving,5,8 meaning of course that the problem of finding a best solution to the alloy Ising problem is not yet a settled problem. There are alloy Hamiltonians that cannot be solved with CE’s. The classical example is the Madelung interaction which, by being long range, cannot be written as a combination of short-range interactions. We will show that there are other circumstances of failure.

The aim of the present paper is to present the results of an in-depth study of CE’s. The CE enthusiasts will find us on the pessimistic side, but we take the point of view that, for a large class of Hamiltonians, CE’s are unavoidable. Thus we study their limitations and possible ways of optimization. Our study is based on the Ga$_{1-x}$Al$_x$N system, on the Madelung Hamiltonian, and on a model of the strain Hamiltonian. The GaAlN system has its own merit because it is a semiconducting system with very large band-gap variation. Thus we take the study of this system as an equally important aim of our work, but the specific results are going to be discussed and published elsewhere.

The Ga$_{1-x}$Al$_x$N is a semiconductor with direct gap at the zone center $\Gamma$ for any concentration $x$. The gaps are not small so that the solutions to the Kohn and Sham equations never run into problems like negative band gaps. The size mismatch between the cations Al and Ga is not large so that the enthalpies of formation of ordered compounds are expected to be small and the relaxation of atomic displacements relatively unimportant. Thus this system is very convenient to study the power and limitations of cluster expansions.

II. CALCULATION METHODS

A. Standard CE and the local-$x$-dependent CE

Our procedure to generate binary configurations follows that of Ref. 9 but extended to general crystalline systems and applied to the wurtzite-hcp case. A list of binary configurations is a necessary step in establishing the parameters of the CE. A standard CE may be written as 1–4,7,10

$$E(\sigma) = \sum_{f} \sum_{n=1}^{N} \frac{1}{V_f} \sum_{k=1}^{V_f} J_f S_n S_{f,k,2} \cdots S_{f,k,V_f},$$

(1)

where $\sigma$ means a binary configuration of atoms (Ga and Al in the present case) and $E$ stands for the enthalpy or any property measurable for all possible configurations. Aside from the enthalpy we are also using a CE for the band gap and for the Madelung and strain models. $f$ means a figure type (empty figure, point, pairs, triangles, etc.) or a collection of $V_f$ vertices, one of which is the site $n$ being summed. $D_f$ is the number of figures of type $f$ per site. There may be many equivalent figures because of the translation and rotational symmetries of the crystal. The product $S_n S_{f,k,2} \cdots S_{f,k,V_f}$ means the product of Ising spins at the $V_f$ vertices of the $k$th...
figure $f$ having one vertex at $n$. In this product, the spin of the second vertex of the figure is $S_{n+k,l}$ and the spin of the $V$th vertex is $S_{n+k,l}$. Equation (1) then says that we must sum over all figures with a vertex at site $n$ and then sum over all sites. Since there are $D_f$ figures of type $f$ per site and those figures have $V_f$ vertices, the sum over the figures $f$ with a vertex at $n$ has $V_f D_f$ terms.

The way the CE is presented in Eq. (1) allows a simple formulation of the local-$x$ CE. Define a local concentration $x_n$ at site $n$ by

$$x_n = \frac{1}{2} + \frac{\overline{S}_n}{2},$$

where $\overline{S}_n$ is a certain average of the spins at the sites neighboring $n$. Assume that the interaction $J_f$ depends on the neighborhood of site $n$. Then

$$J_f(x_n) = \sum_{l=0}^{l_{\max}} J_{f,l}(x_n - 1/2)^l = \sum_{l=0}^{l_{\max}} 2^{-l} J_{f,l} \overline{S}_n^n$$

(2)

and

$$E(\sigma) = \sum_{l=0}^{l_{\max}} 2^{-l} \left\{ \sum_{n=1}^{N} \sum_{f} \frac{1}{V_f} \sum_{k,l} J_{f,l} \overline{S}_n^n S_{f,k,l} \cdots S_{f,k,l'} \right\}.$$

(3)

In the present case the average $\overline{S}_n$ was taken among the 12 nearest neighbors of site $n$. It is convenient to write the enthalpy per site (or any other measurable property) as

$$E(\sigma) = \sum_F J(F) \Pi(F, \sigma),$$

(4)

where

$$\Pi(F, \sigma) = \frac{1}{N} \left\{ \sum_{n=1}^{N} \sum_{f} \frac{1}{V_f} \sum_{k,l} J_{f,l} \overline{S}_n^n S_{f,k,l} \cdots S_{f,k,l'} \right\}.$$

(5)

The symbol $F$ now stands for the pair $\{f,l\}$ and means a generalized interaction.

The generalized interactions we used are based on the traditional pairs, triangles, and tetrahedra listed in Table I. In that table we list the vertex positions of the many figures. The generalized interactions are the products of the spins at the vertices times the powers of 2, 1, and 0 of the local average spin. In Table II we list five different CE types that we considered. In each type, “t,” “tG,” “xG,” “xtG,” and “Full,” we show the powers of the local average spin that were used.

For the no-site interaction $J_0$, which is a configuration-independent constant, we considered it to be dependent on the global (not local) concentration $x$. It is now well known that by making $J_0$ dependent on $x$ one can account for a large part of the elastic energy—that is, the energy resulting from alloying atoms with different radii. In our case we went up to the third power of $x$:

$$J_0(\sigma) = J_{0.0} + J_{0.1} \overline{S} + J_{0.2} \overline{S}^2,$$

(6)

where $\overline{S}$, with no subscript, is the global average of the spin. Observe that we cannot include the first power of $\overline{S}$ because this term would repeat the one-site interaction $J_{1,0}$.

### B. Random arrangement of atoms

When the arrangement is random the site occupations are uncorrelated. This happens at high temperatures, or much larger temperatures than those of ordering or spinodal de-

<table>
<thead>
<tr>
<th>Name</th>
<th>Sites</th>
<th>Name</th>
<th>Sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>$J_0$</td>
<td>(1,2,0,3)</td>
<td>$K_3$</td>
<td>(1,2,0,3) (4, 2, 3, 6) (-2, 2, 3, 6)</td>
</tr>
<tr>
<td>$J_1$</td>
<td>(1,2,0,3) (-2, 2, 3, 6)</td>
<td>$L_3$</td>
<td>(1,2,0,3) (4, 2, 3, 6) (-2, 2, 3, 6)</td>
</tr>
<tr>
<td>$J_2$</td>
<td>(1,2,0,3) (4, 2, 0, 3)</td>
<td>$M_3$</td>
<td>(1,2,0,3) (4, 2, 3, 6) (-2, 2, 3, 6)</td>
</tr>
<tr>
<td>$L_2$</td>
<td>(1,2,0,3) (10, 8, 3, 6)</td>
<td>$N_3$</td>
<td>(1,2,0,3) (4, 2, 3, 6) (-2, 2, 3, 6)</td>
</tr>
<tr>
<td>$M_2$</td>
<td>(1,2,0,3) (1, 2, 3, 3)</td>
<td>$O_3$</td>
<td>(1,2,0,3) (4, 2, 3, 6) (-2, 2, 3, 6)</td>
</tr>
<tr>
<td>$N_2$</td>
<td>(1,2,0,3) (10, 2, 3, 6)</td>
<td>$P_3$</td>
<td>(1,2,0,3) (4, 2, 3, 6) (-2, 2, 3, 6)</td>
</tr>
<tr>
<td>$O_2$</td>
<td>(1,2,0,3) (7, 5, 0, 3)</td>
<td>$Q_3$</td>
<td>(1,2,0,3) (4, 2, 3, 6) (-2, 2, 3, 6)</td>
</tr>
<tr>
<td>$P_2$</td>
<td>(1,2,0,3) (4, 2, 3, 6)</td>
<td>$R_3$</td>
<td>(1,2,0,3) (4, 2, 3, 6) (-2, 2, 3, 6)</td>
</tr>
<tr>
<td>$Q_2$</td>
<td>(1,2,0,3) (7, 2, 0, 3)</td>
<td>$S_3$</td>
<td>(1,2,0,3) (4, 2, 3, 6) (-2, 2, 3, 6)</td>
</tr>
<tr>
<td>$R_2$</td>
<td>(1,2,0,3) (16, 8, 3, 6)</td>
<td>$T_3$</td>
<td>(1,2,0,3) (4, 2, 3, 6) (-2, 2, 3, 6)</td>
</tr>
<tr>
<td>$S_2$</td>
<td>(1,2,0,3)(7, 5, 3, 3)</td>
<td>$U_3$</td>
<td>(1,2,0,3) (4, 2, 3, 6) (-2, 2, 3, 6)</td>
</tr>
<tr>
<td>$T_2$</td>
<td>(1,2,0,3) (10, -4, 3, 6)</td>
<td>$V_3$</td>
<td>(1,2,0,3) (4, 2, 3, 6) (-2, 2, 3, 6)</td>
</tr>
<tr>
<td>$U_2$</td>
<td>(1,2,0,3) (4, 2, 9, 6)</td>
<td>$W_3$</td>
<td>(1,2,0,3) (4, 2, 3, 6) (-2, 2, 3, 6)</td>
</tr>
<tr>
<td>$V_2$</td>
<td>(1,2,0,3) (7, 2, 3, 3)</td>
<td>$K_4$</td>
<td>(1,2,0,3) (4, 2, -3, 6) (-2, -1, 3, 3)</td>
</tr>
<tr>
<td>$W_2$</td>
<td>(1,2,0,3) (7, 2, 3, 3)</td>
<td>$L_4$</td>
<td>(1,2,0,3) (4, 2, -3, 6) (-2, -1, 3, 3)</td>
</tr>
</tbody>
</table>
TABLE II. Different cluster expansions. The CE named t is the standard, including at most 13 pair interactions, 13 triangles, and 3 tetrahedra and has no local spin average dependence. tG is the standard to which we add the terms $a_2 S^2 + a_3 S^3$ burrowed from the old e-G procedure (Ref. 14). xG includes at most pairs and the site interaction with powers 1 and 2 of the local spin average. xtG also includes triangles and tetrahedra, these without local spin average dependence. The CE “Full,” which we never used to fit, includes pair, triangles, and tetrahedra with powers of $S$. tG and xG have equivalent performances, xG having the advantage of establishing a simpler hierarchy of interactions (by the pair sizes), and tG having the advantage of being simpler to program, in a Monte Carlo spin-flip calculation.

<table>
<thead>
<tr>
<th>Name</th>
<th>Powers of $\tilde{S}_{local}$</th>
<th>Powers of $\tilde{S}_{local}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>tG</td>
<td>0    0.2 0.2 0.2 0.2 0.2</td>
<td>tG</td>
</tr>
<tr>
<td>J0</td>
<td>0    0.2 0.2 0.2 0.2 0.2</td>
<td>J0</td>
</tr>
<tr>
<td>K0</td>
<td>0    0.2 0.2 0.2 0.2 0.2</td>
<td>K0</td>
</tr>
<tr>
<td>L0</td>
<td>0    0.2 0.2 0.2 0.2 0.2</td>
<td>L0</td>
</tr>
<tr>
<td>M0</td>
<td>0    0.2 0.2 0.2 0.2 0.2</td>
<td>M0</td>
</tr>
<tr>
<td>N0</td>
<td>0    0.2 0.2 0.2 0.2 0.2</td>
<td>N0</td>
</tr>
<tr>
<td>O2</td>
<td>0    0.2 0.2 0.2 0.2 0.2</td>
<td>O2</td>
</tr>
<tr>
<td>P2</td>
<td>0    0.2 0.2 0.2 0.2 0.2</td>
<td>P2</td>
</tr>
<tr>
<td>Q2</td>
<td>0    0.2 0.2 0.2 0.2 0.2</td>
<td>Q2</td>
</tr>
<tr>
<td>R2</td>
<td>0    0.2 0.2 0.2 0.2 0.2</td>
<td>R2</td>
</tr>
<tr>
<td>S2</td>
<td>0    0.2 0.2 0.2 0.2 0.2</td>
<td>S2</td>
</tr>
<tr>
<td>T2</td>
<td>0    0.2 0.2 0.2 0.2 0.2</td>
<td>T2</td>
</tr>
<tr>
<td>U2</td>
<td>0    0.2 0.2 0.2 0.2 0.2</td>
<td>U2</td>
</tr>
<tr>
<td>V2</td>
<td>0    0.2 0.2 0.2 0.2 0.2</td>
<td>V2</td>
</tr>
</tbody>
</table>

C. First-principles calculations

The total-energy and electronic structure calculations for each configuration to be used in the CE’s were based on the density functional theory (DFT) within the generalized gradient approximation (GGA) for the exchange-correlation potential, proposed by Wang and Perdew. We used the frozen-core projector-augmented-wave (PAW) method as implemented in the “Vienna ab initio simulation package” (VASP) code. The $k$-space integrals were approximated by sums over special-point mesh of the Monkhorst-Pack type within the irreducible part of the Brillouin zone (BZ). The external parameters (lattice vectors) were VASP relaxed only for the extreme wurtzite binaries GaN and AlN. For the other configurations of atoms we used the Vegard law. We made calculations on the (Ga, Al)N configurations of two sets. The smaller set had 63 of the most symmetric configurations of atoms. For this set, the relaxation of the atomic positions (internal degrees of freedom) was carried out in two ways: (i) by the self-consistent program (VASP) itself, by diminishing the Hellmann-Feynman forces until the energy difference between two consecutive changes of atomic positions was $\leq 10^{-4}$, and (ii) by a force field (FF) to be described later. For the larger set of 327 not so symmetric configurations, the relaxation of the atomic positions was made with the FF only. In this case the self-consistent program was used at fixed atomic positions to minimize the electronic energy.

D. Force field for the Ga$_{1-x}$Al$_x$N system

To generate a very large set of enthalpies and band gaps, instead of calculating full VASP relaxed configurations of N, Ga, and Al in the wurtzite lattice, we used a force field that could give good, though not perfect, atomic positions. The whole procedure was the following. (i) Full-relaxation VASP calculation of wurtzite GaN and AlN. (ii) Use of linearity (Vegard) of the lattice parameters, both perpendicular and parallel to the c axis, for any configuration Ga$_{1-x}$Al$_x$N. (iii) Relaxation of the internal degrees of freedom (atomic positions $\vec{r}_i$ and $\vec{r}_j$) by minimizing the following FF energy functional:

$$\sum_{i<j} A_{ij} (|\vec{r}_i - \vec{r}_j|^2 - R_{ij}^2)^2.$$  

The parameters $R_{ij}$ are the following. $R_{Ga,Ga}$ and $R_{Ga,N}$ are neighbor distances taken from the VASP results for GaN. $R_{Ga,Al}$ and $R_{Al,N}$ were taken from the VASP results for AlN. $R_{Ga,Al}$ was chosen as an average between $R_{Ga,Ga}$ and $R_{Al,Al}$. The atomic positions have just a negligible dependence on the amplitudes $A$ in Eq. (8), so we set $A_{Ga,Ga} = A_{Al,Al} = 100$, $A_{N,N} = 0$, and $A_{Ga,N} = A_{Al,N} = 1$. With this choice of parameters, we make the cation lattice of Ga and Al almost independent of the N positions. After establishing the cation positions the N move inside the cation tetrahedra so as to minimize the N energy terms.

In Fig. 1 we plot the enthalpies of formation for the configurations $\sigma$,$$

\Delta E(\sigma) = E(\sigma) - (1 - x) E(Ga) - x E(Al),$$

calculated with relaxation of the atomic positions (internal degrees of freedom $\vec{r}_i$) by means of the force field and by means of VASP itself. Clearly, the difference between the two relaxations is much smaller than the dispersion of the many enthalpies. Only for the high energy configurations, those that have much to lose on relaxation, does the FF relaxation seem to be insufficient. The figure gives us an assurance that the study we are making on the CE is on a system that is very near a real system.
FIG. 1. Enthalpies of formation of 63 configurations calculated from first principles with relaxation of the atomic positions by VASP and by the force field described in the text. The 63 configurations were those of highest symmetry. The hcp space group, not the wurtzite space group, was used to define the interactions (figures) and configurations. The configuration labeled as 2+ is B$_{1a}$ tungsten carbide. 136+ are configurations with $\mathbf{A}_{1,2}=2\mathbf{a}_{1,2}$ whose space groups are 4 times smaller than the hcp space group. 5+ is a configuration with $\mathbf{A}_1=2\mathbf{a}_1$ and other unit vectors equal to those of hcp and whose space group is 6 times smaller.

III. PROCEDURES

To test a CE fit to first-principles results of a set of configurations one needs a larger set in which the CE is tested, aside from the smaller set (the training set) to which the CE is fit. Thus we made first-principles calculations, using VASP, on two sets. The smaller set had 63 configurations and was calculated in two different ways: (i) the atomic positions were relaxed by VASP itself, and (ii) the atomic positions were relaxed by a force field (FF). As already said before, in both cases the lattice vectors were determined by the Vegard interpolation of GaN and AlN. The 63 configurations were those whose space-group size was at least 1/8 of the GaN and AlN space groups. The first-principles calculations will always deal with the configurations with larger symmetry, at least not a poor translation symmetry due to a very large unit cell. The larger set had 327 configurations and included the configurations of the smaller set and all configurations whose unit cell had at most 8 cations. The relaxation of the atomic positions of the larger set was made with the FF only. The larger set was used to assess the quality of the CE’s whose parameters were determined from the VASP results of the smaller set red (training set) with FF relaxation.

Two comments are in order. First we must stress that, in defining the CE parameters by means of the smaller set of configurations, the relaxation of that set was made by means of the force field, not by VASP, as for the relaxation in the larger set. Second, it is frequently argued that the best way to find the CE parameters is through a small set of configurations that also includes those of low symmetry. Now, our small set was made of configurations with the largest symmetries, yet the number 63 is so large that most of those configurations have poor symmetry. In studying the quality of CE’s we definitely did not want to bias the CE generation by choosing small-set configurations by any particular criterion not transferable to other crystalline systems. To make this point clearer we are postponing to Appendix A a discussion on the possible inclusion of the “special quasirandom structures” (SQS’s) in the training set. These configurations are among those with the poorest symmetry, and yet their inclusion in the 63 configurations training set would not modify the CE’s in any important way.

A. Fitting the interaction parameters to the smaller set of configurations

Reference 10 gives a very understandable way to make the fit. For a given set of configurations, for which one has first-principles results, and for a given set of interactions, their values are to be determined by minimizing the root-mean-squared (rms) error of the fitted with respect to the true value. Of course this recipe does not answer the question of how to choose the interactions. To answer this very important question, Ref. 10 uses a figure of merit that corresponds to the predictive power of the set of interactions. The idea is the following. Let $\sigma$ be a configuration of the set, let $e(\sigma)$ be the first-principles value (total energy or band gap of GaAlN, or Madelung or strain model energies) for the configuration, let $F$ be an interaction (figure) of the set, $J(F)$ its value, and let

$$E(\sigma) = \sum_F J(F)\Pi(F,\sigma)$$

be the cluster expansion approximation to the true value $e(\sigma)$. Here $\Pi(F,\sigma)$ have received different names like “correlation” or “product of spins,” names that are inadequate to the coefficients of more general CE’s such as the $x$-dependent of Ref. 8. We shall refer to these coefficients simply as $\Pi$.

If the set of interactions and the set of configurations are given, the interaction values $J(F)$ should be chosen so to minimize the rms error:

$$\text{rms}^2 = \frac{1}{N} \sum_{\sigma=1}^{N} [e(\sigma) - E(\sigma)]^2 = \text{min}.$$  \hspace{1cm} (9)

This minimization brings no information on the predictive power of the set of interactions. To know its predictive power we consider the set of configurations with one of them excluded—say, configuration $\omega$. With this exclusion we re-calculate the values $J(F)$, again using Eq. (9), and obtain the approximation $\hat{E}(\omega)$ to the first-principles-calculated value corresponding to the excluded configuration. Following Ref. 10 we define the “cross-validation” (CV) figure of merit as

$$\text{CV}^2 = \frac{1}{N} \sum_{\omega=1}^{N} [e(\omega) - \hat{E}(\omega)]^2; \hspace{1cm} (10)$$

in other words, we sum squared errors for each configuration when it is excluded from the set. As a practical way to calculate CV one proves the relation
subtraction we looked for the interactions to be added or subtracted that resulted in the smallest CV. Thus our search for a global CV minimum was based on six independent sequences: scanning the CV values by increasing the number of interactions in steps of 1, 2, and 3, and by decreasing the number of interactions. Of course we chose the set of interactions leading to the minimum CV value among the minima for each of the 6 scans. It is difficult to compare this procedure with the genetic algorithm but the different scans lead to rather flat curves around the minima that having the true global minimum is not important. Figure 2 shows an exceptional situation when the very CV criterion is failing (the case of a xtG CE in a scan of increasing the number of interactions by steps of 2) because it leads to a very low CV minimum for the 63-configuration set but a 4-times larger rms error for the large 327-configuration set.

B. Long-range interactions

One of our purposes is to study to what extent can the long-range interactions be written as a CE. We recognize long-range interactions of two kinds. The first is, of course, the Madelung (electrostatic) interaction

\[ E(\text{Madelung}) = A \sum_{i,j} \frac{S_i S_j}{|r_i - r_j|}, \]

which is very well known and can be calculated by the famous Ewald method. Of course the Hamiltonian above is already written as CE’s of pairs. Then the question is how well this Hamiltonian can be reduced to a CE with short-range interactions, pairs, or other clusters.

Not as known is that part of the elastic interaction that results from associating different spins ±1 with different atomic sizes. In Appendix B we show that, among all elastic interactions, the one with longer range writes as

\[ E(\text{strain}) = B \sum_{i,j} \frac{S_i S_j}{|r_i - r_j|^3}. \]

Again this is a pair-CE Hamiltonian that needs be to written in terms of short-range interactions. Though the pair interaction falls down with distance much faster than Madelung, \( r^{-3} \) is still long range as one verifies by integrating in a sphere of radius \( R \). As \( R \rightarrow \infty \) there is a logarithmic divergence.

It must be said that our study of these two Hamiltonians is quite independent of the Ga\textsubscript{1-x}Al\textsubscript{x}N case, which is an alloy of isovalent atoms. We are searching for the \( 1/r^2 \) pair interaction that can be reduced to a short-range CE.

IV. RESULTS AND DISCUSSION

A. Optimal CE’s and their predictive power

In Table I we present the interactions that we used. In the case of the local-\( x \)-CE, the local average \( \bar{S}_i \) at the site \( i \) was made equal to the average of the spins of the 12 first neighbors. Aside from the sites defining the interaction, one must also specify the power \( l \) of \( \bar{S}_i \) (see Sec. II A). In the case of the standard CE that power is always zero. In Table II we tabulate the different CE’s that we considered. They differ by...
This comparison is in Figs. 3–6, for the enthalpy and band principles results and compared them to what VASp gave. With those interactions we “predicted” the first-configurations to establish the interaction values at the minimum strain energies of GaAlN and for wurtzite models of the Madelung and Si the chosen figures and by the chosen powers of the local spin and minimum CV the CE were determined with the smaller set of 63 configurations [Eq. (B1)].

As explained earlier, we used the smaller set of 63 configurations to establish the interaction values at the minimum CV. With those interactions we “predicted” the first-principles results and compared them to what VASP gave. This comparison is in Figs. 3–6, for the enthalpy and band gap of GaAlN and for wurtzite models of the Madelung and strain energies [Eq. (B1)].

Figure 3 compares enthalpies of formation, CE’s and first-principles (VASP) calculated, for the 327 configurations. A perfect CE would produce a figure with all crosses and squares along the diagonal. The CE we used in all fits was that giving the smallest CV for the smaller set of 63 configurations and minimum CV (squares). In the present case the minimum occurred for a xtG CE (see Table II) with 14 interactions. The parameters of the CE were used to “predict” the Madelung energy of the 327 configurations of the larger set (crosses). As expected, the cluster expansions perform poorly for the Madelung energy because it is a long-range interaction. The large values of the Madelung energies correspond to the [m,n] superlattices along the c axis. These are configurations marked as 202+ (a [4,4] superlattice), 195+ ([5,3]), and 37+ ([3,3]).

Figure 4 compares the band gaps at the zone center (Γ) CE and VASP calculated. We are plotting the quantity

$$\Delta \text{gap}(\sigma) = \text{gap}(\sigma) - (1 - x) \text{gap}(\text{GaN}) - x \text{gap}(\text{AlN}).$$

Again one would like to have squares and crosses falling at the diagonal, but that turned out to be impossible. Figure 5 compares the Madelung energy calculated exactly and by the CE. Of course the CE is wholly inadequate to this long-range interaction. Fortunately we are alloying isoelectronic atoms
The result is represented in Fig. 6 to be large. Unlike the Madelung interaction, the strain interaction can be described by a CE. The parameters of the CE were determined with the smaller set of 63 configurations and minimum CV (squares). In the present case the minimum occurred for a 4xG CE (see Table II) with 14 interactions. The parameters of the CE were used to “predict” the strain energy of the 327 configurations of the larger set (crosses). Though the strain interaction may be considered long range, the cluster expansion works very well and most points do fall on the diagonal. The large values of the strain energies correspond to the \([m,n]\) superlattices along the \(c\) axis, as in the case of the Madelung energy. These are configurations marked as 202+ (a [4,4] superlattice), 195+ ([5,3]), and 37+ ([3,3]). The names of the worst fitting configurations are marked (57+, 61+, 45; see Fig. 7).

as Ga and Al so that a Madelung contribution is not expected to be large.

Unlike the Madelung interaction, the strain interaction can be described by a CE. The result is represented in Fig. 6 where we notice that the many points stay closer to the diagonal than for the enthalpy and gap. As in the case of the Madelung interaction, the large interactions are for the long-period superlattices 37+, 195+, and 202+, but they stay at the diagonal of the figure. Curiously, the largest deviations from the diagonal are for the same configurations as in the cases of enthalpy and gap. As in the case of the band gap, the superlattice will tend to have the gap of the extreme with the smallest value, not any sort of average between the gaps of the extremes.

FIG. 6. Strain energy, exact against calculated with a minimum CV cluster expansion. The parameters of the CE were determined with the smaller set of 63 configurations and minimum CV (squares). In the present case the minimum occurred for a 4xG CE (see Table II) with 14 interactions. The parameters of the CE were used to “predict” the strain energy of the 327 configurations of the larger set (crosses). Though the strain interaction may be considered long range, the cluster expansion works very well and most points do fall on the diagonal. The large values of the strain energies correspond to the \([m,n]\) superlattices along the \(c\) axis, as in the case of the Madelung energy. These are configurations marked as 202+ (a [4,4] superlattice), 195+ ([5,3]), and 37+ ([3,3]). The names of the worst fitting configurations are marked (57+, 61+, 45; see Fig. 7).

FIG. 7. Configuration 57+, for which the CE fits are the worst. The light gray balls represent Ga atoms, and the dark balls are Al. The figure pictures two planes perpendicular to the \(c\) axis. The next plane repeats the first because the configuration unit vector \(\vec{A}_1=\vec{a}_3\), the unit vector of the Bravais hexagonal lattice. This configuration 57+ is a [4,4] superlattice with a lower symmetry than the [4,4] superlattice along the \(c\) axis. The configuration 45, also a bad CE fit, is a [5,3] superlattice along the same direction. The configuration 61+, another with a bad fit, is also [4,4] but differs from the figure above because the light and the dark balls are displaced by one plane, which makes the interplane distances different from those of the figure above.

GaN and AlN, one extreme at each region of the superlattice. In the case of the band gap, the superlattice will tend to have the gap of the extreme with the smallest value, not any sort of average between the gaps of the extremes.

### B. GaAlN alloy system

Once the interactions \(J_{ij}\) are known, we can use Eq. (7) to calculate the average enthalpy of the random alloy at any composition \(x\):

\[
H(x) = \langle E(\sigma) \rangle - xE(\text{Al}) + (1-x)E(\text{Ga}),
\]

from which we obtain the Bragg-Williams free energy per site,

\[
G = H(x) + kT x \ln(x) + kT (1-x) \ln(1-x).
\]

The estimate of the spinodal decomposition temperature is obtained from the second derivative

\[
\frac{\partial^2 G}{\partial x^2} = 0,
\]

which gave 270 K, at \(x=0.5\), for the GaAlN system. This means that, at the practical temperatures of growth, AlN and GaN are very miscible. This is a common feature of all semiconductor alloy systems whose pure binaries have small lattice mismatch.

### V. CONCLUSIONS AND SUMMARY

Our study of the many CE leads us to the following conclusions: (i) CE’s are reliable except for the long-period superlattices. (ii) The CV criterion of quality remains as the only one available, but we found an instance where it was misleading. (iii) The Madelung energy cannot be simulated by any short-range CE. (iv) The strain energy is very well described by a CE, and we found that it also serves as a guide to choose CE’s for enthalpy and gap. Of course it will...
also fail for the long period superlattices. (v) Among the many proposals for choosing figures (interactions), the \( x \)-dependent interactions are very convenient because they provide a simple hierarchy, that of the pair lengths. On the other hand, programing \( x \)-dependent interactions is complicated. (vi) The optimal set of interactions is the largest linear independent set. (vii) Adding terms like \( x^2 \) and \( x^3 \), of the \( e-G \) procedure, can improve the standard CE considerably.

Our study on the GaAlN system shows that it is a system with very high miscibility. Practically it should form solid solutions at the growth temperatures. Large gap deviations from linearity can be obtained by preparing long-period superlattices.

ACKNOWLEDGMENTS
This work was supported by the Brazilian funding agencies FAPESP and CNPq.

APPENDIX A: LOW-SYMMETRY CONFIGURATIONS: THE SQS
The issue here is the inclusion of low-symmetry configurations, such as the SQS, in the training set. First we must say that any training set will be made of not-too-low-symmetry configurations because the unit cells are never too large. Yet between our space group \( \geq hcp \) training set configurations and the \( hcp \) of some of our SQS there is an important symmetry reduction that must be investigated.

In Fig. 9 we plot CE-calculated enthalpies against the VASP calculated for the training set configurations and for some good SQS candidates. The SQS candidates had unit cells with at most eight cations and small (minimum and close to it) values of

\[
\frac{1}{13 \text{ pairs}} \sum (\bar{S}_i \bar{S}_j - \bar{S}_i \bar{S}_j)^2.
\]

Some features are outstanding: (i) if the SQS truly represented the random alloy, all would have the same first-principles-calculated enthalpy and that is not the case; (ii) the points corresponding to the SQS candidates have a deviation from the ideal diagonal that is only slightly larger than that of the training set points themselves. This means that the 63 configurations training set is able to represent the SQS fairly well. The SQS are in no way exceptional as are the long-period superlattices. Probably with their inclusion in the

![FIG. 8. CE band gaps, corrected with a linear term to account for the experimental values, versus the composition \( x \). Observe that the long-period superlattices may have band gaps quite smaller than the zero bowing line (linear interpolation). Configurations 194+, 194, and 202+ are superlattices along the \( c \) axis. Configuration 2+ is of type tungsten carbide.](image)

![FIG. 9. CE and first-principles enthalpies. Ideally all points should fall along the diagonal. The SQS points deviate from the diagonal slightly more than the 63 configurations training set points, meaning that they are well described by the CE.](image)
where the displacements \( u \) per atom, each having “strength” \( B \) typical of its species. We assume that the crystal is made of these singularities, one per atom, each having “strength” \( B \) typical of its species. With the displacements \( u \) we obtain the strain field by derivation and the elastic energy by squaring the components of strain and integrating. Now we are interested only in the energy differences between a certain configuration of atoms and a random distribution with the same concentration. Thus, we obtain the strain field by derivation.

\[
\tilde{u}(\vec{r}) = -B \nabla \left( \frac{1}{|\vec{r} - \vec{r}_0|} \right).
\]

We assume that the crystal is made of these singularities, one per atom, each having “strength” \( B \) typical of its species. The energy per atom is then

\[
E = \frac{C}{N} \sum_i \left( \frac{B B_i}{|\vec{r}_i - \vec{r}_j|^3} \right),
\]

where \( C \) is a constant and \( N \) is the number of sites. This is similar to the Madelung energy (electrostatic) but the \( 1/r \) dependence of the latter becomes \( 1/r^3 \). Though \( 1/r^3 \) falls faster with distance than \( 1/r \), it is still a long-range interaction as it has a logarithm divergence in an all-space volume integration.

The way to deal with the sums in Eq. (B1) is similar to the Ewald method (a sum in the reciprocal space plus a sum in the direct space). Letting \( S_i \) be the spin at the site \( i \) and letting \( \tilde{S} \) be the average spin, the energy that interests us is

\[
E = \frac{D}{N} \sum_{i,j} \left( S_i - \tilde{S} \right) \left( S_j - \tilde{S} \right) \frac{1}{|\vec{r}_i - \vec{r}_j|^3},
\]

and the Ewald method gives

\[
\sum_i \frac{1}{|\vec{r}_i - \vec{r}|^3} = \frac{4\pi}{N\Omega} \ln G + \frac{2\pi}{N\Omega} \sum_{g \neq 0} e^{\tilde{g} \cdot \vec{r}} \left[ \text{Ei} \left( 1, \frac{g^2}{4G^2} \right) - \exp \left( -\frac{g^2}{4G^2} \right) \right] + \frac{2}{\sqrt{\pi}} \sum_i \sqrt{\pi} \operatorname{erfc}(\vec{r} - \vec{r}_i G) + 2G|\vec{r} - \vec{r}_i| \left( G^2|\vec{r} - \vec{r}_i|^2 + 1 \right) e^{-\frac{1}{2}G^2|\vec{r} - \vec{r}_i|^2},
\]

and

\[
\sum_{i \neq j} \frac{1}{|\vec{r}_i - \vec{r}_j|^3} = \frac{4\pi}{N\Omega} \ln G + \frac{2\pi}{N\Omega} \sum_{g \neq 0} \left[ \text{Ei} \left( 1, \frac{g^2}{4G^2} \right) - \exp \left( -\frac{g^2}{4G^2} \right) \right] + \frac{2G^3}{3\sqrt{\pi}} + \frac{2}{\sqrt{\pi}} \sum_{i \neq 0} \sqrt{\pi} \operatorname{erfc}(IG) + 2GI(G^2I^2 + 1)e^{-G^2I^2},
\]

where

\[
\text{Ei}(n, \chi) = \int_1^n \frac{e^{-\chi t}}{t} dt = \text{the exponential integral}.
\]


15 We maintained the same parameters for all calculations (files INCAR and KPOINTS) to make the results for the many configurations comparable.

16 Among the configurations with space group 8 only two would not converge without modifying the set of parameters controlling the VASP convergence. We decided not to modify the parameters to maintain the configurations all calculated consistently.


