Exploring the configurational space of binary alloys: Different sampling for different cell shapes

G. Trimarchi, P. Graf, and A. Zunger*
National Renewable Energy Laboratory, Golden, Colorado 80401, USA

(Received 12 March 2006; published 14 July 2006)

In many areas of alloy theory, such as determination of the $T=0$ ground state structures or calculation of finite-$T$ alloy thermodynamics, one needs to enumerate and evaluate the $\sim 2^N$ configurations $\sigma$ created by different substitutions of atoms A and B on the N sites of a unit cell. These configurations consist of $M_{\text{ICS}}$ “inequivalent cell shapes” (ICS’s), each having $M_{\text{SSS}}$ “same-shape structures” (SSS’s). Exhaustive evaluation approaches attempt to compute the physical properties $P(\sigma)$ of all SSS’s belonging to all ICS’s. “Inverse band structure” approaches sample the physical properties of all SSS’s belonging to a single inequivalent cell shape. We show that the number $M_{\text{ICS}}$ of ICS’s rises only as $BN^{n}$, whereas the total number of SSS’s scales as $AE^{8n}$. Thus, one can enumerate the former (i.e., calculate all) and only sample the latter (i.e., calculate but a few).

Indeed, we show here that it is possible to span the full configurational space efficiently by sampling all SSS’s (using a genetic algorithm) and repeating this by explicit evaluation for all ICS’s. This is demonstrated for the problem of ground state search of a generalized cluster expansion for the Au-Pd and Mo-Ta alloys constructed from first-principles-total-energy calculations. This approach enables the search of much larger spaces than hitherto possible. This is illustrated here for the $2^{32}$ alloy configurations relative to the previously possible $2^{20}$.

DOI: 10.1103/PhysRevB.74.014204

PACS number(s): 61.50.−f, 61.66.−f

I. INTRODUCTION: GENERATING A CONFIGURATIONAL SPACE $\sigma$ AND SURVEYING PHYSICAL PROPERTIES $P(\sigma)$ IN THAT SPACE

Most structural optimization issues in solids can be classified as being either positional refinement or configurational search problems. In positional refinement problems one is given the global topology, composition, and size of a system (molecule, solid) and needs to locally relax the atomic positions to reach the equilibrium structure.1 In configurational search problems2 one is exploring the different assignments of atoms to the lattice sites in unit cells of various shapes and sizes, e.g., different structure-types such as $L1_0$, $DO_{22}$, etc. While ideally it would be desirable to have a single optimization strategy for both types of search problems, it is often convenient to study positional refinement problems by continuous-motion techniques such as molecular dynamics,3 whereas configurational search problems are best studied by discrete sampling approaches2 such as Monte Carlo3 or direct enumeration3,4 (followed by local atomic relaxation). Here, we discuss configurational search problems.

Indeed, a binary $A_1$,$B_2$ alloy of composition $x$ can appear in as many as $\sim 2^N$ spatial lattice configurations describing the different realizations obtained by occupying each of the $N$ sites of a repeated unit cell of a fixed lattice by atom $A$ or $B$. These $\sim 2^N$ configurations contain ordered structures, superlattices, isolated and clustered impurities etc., and represent all possible equilibrium and nonequilibrium structures. In principle, each configuration $\sigma$ could have its own distinct physical property $P(\sigma)$, such as the total-energy, strain, band gap or Curie temperature. The need to enumerate and define the $2^N$ lattice configurations $\sigma$ appears in many areas of alloy theory. For example, when one searches the $T=0$ ground state configuration one must consider the energy $E(\sigma)$ of all $2^N$ configurations, thus seeking the lowest “tie-line.”2,4 When one needs to describe a random alloy, a configurational average $\langle P(\sigma) \rangle$ over all $P(\sigma)$ is needed.2 In designing materials with a target property $P_{\text{target}}$ via the inverse band structure5 (IBS) method one minimizes the objective function $O(\sigma) = \langle P(\sigma) - P_{\text{target}} \rangle^2$ by varying the configuration variables $\sigma = \{S_1, S_2, \ldots, S_N\}$ (where the site-occupation variables $S_i$ take values +1 or −1, respectively, when they represent an atom of type A or type B). In all of the above examples one needs to calculate all $P(\sigma)$. There are approaches in alloy theory which circumvent the need to define and enumerate all configurations. For example, traditional first-principles ground state searches6–10 look for the lowest total energy among a small chosen list of only $\sim 10$ candidate configurations. However, the potential of missing lower energy configurations is large. In the coherent potential approximation11 (CPA) to an alloy, the configuration averages are calculated without the need to enumerate all $\sigma$’s, since all atomic configurations $\sigma$ are replaced by a virtual medium in which each of the A atoms (and separately, each of the B atoms) experiences an identical lattice environment. This drastic approximation misses the correct electrostatic energy of the alloy.12,13 However, it is generally important to find methods which define and enumerate the various configurations in the $2^N$ space, without approximations, thus enabling calculations of $P(\sigma)$. In this paper we will describe a simple procedure for generating the configurational space and strategies to search it for specific $P(\sigma)$. This will allow us to perform ground state searches or to find structures that have target properties.

We will see in this paper (Fig. 1) that the set of configurations that have $N$ lattice sites per unit cell, can be divided, considering the translational symmetry, into groups $I=1,\ldots,N_{\text{ICS}}$, each having a distinct “inequivalent cell shape” (ICS), where $N_{\text{ICS}}$ is the total number of ICS’s containing $N$ atoms. Within each such group I there are a number of compositions $A_n B_{N-n}$ (where $n=1,\ldots,N-1$), and each composition has a set of “same-shape structures” (SSS’s). There are two general search strategies for surveying $\sigma$: (i) exhaustive evaluation, where $P(\sigma)$ is calculated for each and every $\sigma$, i.e., all ICS’s and all SSS’s, and (ii) sampling, where only a few $\sigma$’s are calculated explicitly. The exhaustive evaluation
FIG. 1. Scheme that summarizes the internal structure of the configurational space of the binary alloy that contains a maximum of $Q$ lattice sites per unit cell: (a) The set of configurations containing $N$ atoms ($N=2, \ldots, Q$) breaks into (b) $N_{ICS}(N)$ subsets, each one characterized by an inequivalent cell shape (ICS). Each of these ICS’s contains (c) $N-1$ subsets, each one characterized by a concentration $x_n=n/N$ (d) For each $I$ and $x_n$ there are $N_{SSS}(N, x_n)$ same-shape-structures (SSS’s).

The remainder of the paper is organized as follows. In Sec. II, we discuss the structure of the configurational space of the binary alloys, showing the hierarchy into which the crystal configurations are organized. We also present the algorithm to determine the configurations that are contained in the space. In Sec. III, we discuss the direct enumeration based “exhaustive evaluation” and the “sampling search,” and in Sec. IV, we combine direct enumeration and sampling procedures to introduce the different sampling for different (cell) shape (DSDS) search method. In Sec. V, we apply this method to the search for the ground state configurations in the metallic alloy Au-Pd and Mo-Ta. In this case we use as $P(\sigma)$ the cluster expansion (CE) expression for the formation energy $\Delta H_{CE}(\sigma)$. We find that searching by DSDS structures with up to $N=20$ atoms gives the same result as exhaustive evaluation on the same configurational space. The application of DSDS for searching structures with up to $N=32$ atoms identifies for Au-Pd two new ground states that were not detected by the exhaustive evaluation performed up to 20 atoms. Interestingly, one of the structures discovered is located in the Pd poor region of the phase diagram and is a superlattice that represents an example of an “adaptive” structure (see Ref. 23) showing a microscopic geometry that had not been unveiled by the exhaustive evaluation restricted to structures of a smaller size unit cell. For Mo-Ta the DSDS search of structures with up to $N=32$ atoms does not identify new ground states with respect to those already obtained by exhaustive evaluation. The final section is devoted to our conclusions.

II. GENERATION OF THE BINARY ALLOY CONFIGURATIONAL SPACE

We consider a binary alloy of the type $A_{1-x}B_x$ described by an underlying fcc or bcc lattice structure. In Figs. 1(a)–1(d) we illustrate the partitioning of the configurational space. Here, (a) $Q$ is the maximum number of atoms per cell, so the configurational space is defined as the set of all the alloy structures, with configuration $\sigma=\{S_1, S_2, \ldots, S_N\}$ and having $N=2, \ldots, Q$ atoms per cell. (b) Considering the structures with $N$ lattice sites, there are $N_{ICS}(N)$ inequivalent cell shapes (ICS’s). Each ICS is specified by a triad of lattice vectors $\{a_1, a_2, a_3\}$ and defines the decoration of the lattice according to the translational symmetry. (c) Each ICS, for example the $I_{th}$ containing $N$ lattice sites, encompasses a set of $N-1$ concentration $x_n=n/N$. (d) Each ICS and concentration have associated a group of same-shape structures (SSS’s).

A. Determination of the “inequivalent cell shapes” containing $N$ lattice sites

Here, we summarize the procedure (introduced by Ferrari, Wei, and Zunger) for the construction of the configurational space composed of all configurations whose unit cell contains $N(\leq Q)$ atoms. This algorithm is systematic and exhaustive as it allows the enumeration of all nonequivalent crystal structures without repetitions. We seek a procedure constructing for a binary alloy the set of unit cells containing $N$ lattice sites and represented by triads of unit cell vectors $\{a_1, a_2, a_3\}$. The underlying topology of the alloy lattice is defined by the triad of unit vectors $A_1, A_2, A_3$: here we will consider the case of fcc and bcc lattices. Choosing the lattice parameter $a=2$ the unit vectors of the fcc lattice are $A_1=(0,1,1), A_2=(1,0,1)$, and $A_3=(1,1,0)$. For the bcc lattice, they are $A_1=(-1,1,1), A_2=(1,-1,1)$, and $A_3=(1,1,-1)$. The vectors $a_i$ defining the unit cell of the alloy are integer combinations of $A_i$ and take the form of triplets of integers, $(l_i, m_i, n_i)$. To enumerate the triads $\{a_i\}$ we proceed as follows:
In the fcc case we require for the vectors \( \mathbf{a}_i \) (for the bcc lattices multiply \( N \) by 2 here and in all the formulas of this section)

\[
\mathbf{a}_1 \times \mathbf{a}_2 \cdot \mathbf{a}_3 = 2N
\]  

(1)

and, with no loss of generality,

\[
a_1^2 \geq a_2^2 \geq a_3^2.
\]  

(2)

Since adding to or subtracting from \( \mathbf{a}_2 \) a multiple of \( \mathbf{a}_1 \) does not change the lattice, we require

\[
(a_2 \pm a_1)^2 \geq a_2^2 \quad \text{or} \quad a_1^2 = 2a_1 \cdot a_2 \geq 0.
\]  

(3)

One may also add any integer combination of \( \mathbf{a}_1 \) and \( \mathbf{a}_2 \) to \( \mathbf{a}_3 \) without changing the lattice, so that we also require

\[
\mathbf{a}_3 = \mathbf{a}_1 + \alpha \mathbf{a}_2 + \gamma \mathbf{a}_1 \times \mathbf{a}_2,
\]  

(4)

where for \( \gamma, \alpha \), and \( \beta \) the following relations are valid:

\[
\gamma = \frac{2N}{(\mathbf{a}_1 \times \mathbf{a}_2)^2}
\]  

(5)

and

\[
-0.5 < \alpha \leq 0.5
\]  

(6)

\[
-0.5 < \beta \leq 0.5.
\]  

(7)

From these relations we obtain the following relations for \( a_1^2 \), \( a_2^2 \), and \( a_3^2 \):

\[
a_1^2 \leq (16N^2)^{1/3},
\]  

(8)

\[
a_2^2 \leq \frac{a_1^2}{3} + \left[ \frac{a_1^4}{9} + \frac{16N^2}{3a_1^2} \right]^{1/2},
\]  

(9)

\[
a_3^2 \leq \frac{4N^2}{(\mathbf{a}_1 \times \mathbf{a}_2)^2} + \frac{a_1^2 + a_2^2 + 2a_1 \cdot a_2}{4}.
\]  

(10)

We use the relations (3)–(10) reported above to determine the triad of vectors \( \{\mathbf{a}_i\} \). For \( \mathbf{a}_1 \), due to the rotational symmetry of the cubic Bravais lattice, we can require that

\[
l_1 \geq m_1 \geq n_1 \geq 0,
\]  

(11)

moreover, we choose \( (l_1, m_1, n_1) \) so that it satisfies Eq. (8).

We choose \( \mathbf{a}_2 = (l_2, m_2, n_2) \) so that Eqs. (3) and (9) are satisfied. Finally, choose \( \mathbf{a}_3 = (l_3, m_3, n_3) \) so that Eqs. (10) and (4)–(7) are satisfied.

The procedure just described may generate cells that are equivalent by symmetry. So to avoid repetitions we verify they are not related by point group operations of the Bravais lattice. Formally, taking two cells, represented by the triads \( \{\mathbf{b}_p\} \) and \( \{\mathbf{a}_q\} \), we verify that there are not point group operations \( R \) of the Bravais lattice that satisfy the relation

\[
\mathbf{b}_p = \sum_q C_{pq} \mathbf{R}_q \mathbf{a}_q,
\]  

(12)

where \( C_{pq} \) is a matrix of integer numbers. The cells containing \( N \) lattice sites, found using the procedure described by the Eqs. (2)–(12), form the set of all ICS’s containing \( N \) lattice sites. Figure 2(a) shows the number \( N_{ICS}(N) \) of ICS’s as a function of \( N \) as obtained by applying the procedure described in the preceding section. This figure shows that the number of ICS’s obtained by direct enumeration is fit well with the function \( BN^\alpha \) where \( \alpha = 3/2 \).

**B. Number of same-shape structures**

Given an ICS with \( N \) basis sites, the stoichiometries of a binary alloy will be \( A_xB_{1-x} \), where the integer \( n \) goes from 1 to \( N−1 \). For the \( j \)th ICS and for composition \( A_xB_{1-x} \), we find that the number of SSS’s, not taking into account symmetry, is

\[
N_{SSS}(N, x_n) = \binom{N}{n},
\]  

(13)

and this number does not depend on the particular ICS considered. Symmetry arguments can reduce this number. For a
given composition \( A_nB_{N-n} \), the symmetry-reduced number \( N_{SSS}^{(l)}(N,x_n) \) of SSS’s will be lower than \( \binom{N}{n} \) and will depend on the \( l \)th cell shape:

\[
N_{SSS}^{(l)}(N,x_n) = D_{N,x_n}^{(l)} \binom{N}{n},
\]

(14)

where \( D_{N,x_n}^{(l)} \leq 1 \) represents the reduction factor due to the application of the symmetry operations. Given the number \( N \) of atoms per cell and the composition \( x_n \), the symmetry reduction factor \( D_{N,x_n}^{(l)} \) will depend on the shape of the \( l \)th ICS.

It is interesting to calculate for the \( l \)th ICS containing \( N \) atoms the total number of SSS’s obtained summing over all the allowed compositions. In case symmetry is not used, the total number of SSS’s is

\[
T_{SSS}(N) = \sum_{n=1}^{N-1} \binom{N}{n} = 2^N - 2
\]

(15)

for any cell shape considered. When symmetry is used, we indicate the total number of SSS with \( T_{SSS}^{(l)}(N) \) and it is given by the relation

\[
T_{SSS}^{(l)}(N) = \sum_{n=1}^{N-1} N_{SSS}^{(l)}(N,x_n) = \sum_{n=1}^{N-1} D_{N,x_n}^{(l)} \binom{N}{n}
\]

(16)

and depends on the shape of the \( l \)th ICS. The value of \( T_{SSS}^{(l)}(N) \) has been evaluated by directly scanning all the structures with up to 20 atoms per cell. In particular we considered the cell shape for which the total number of SSS’s is 

**III. SEARCHING THE BINARY ALLOY CONFIGURATIONAL SPACE: EXHAUSTIVE EVALUATION VS SAMPLING**

**A. Exhaustive evaluation of the configurational space for \( N \leq Q \)**

In the exhaustive evaluation (EE) approach the physical property \( P(\sigma) \) is computed for all ICS’s, concentrations \( x_n \), and the corresponding SSS’s (see Fig. 1) that contain up to \( Q \) lattice sites. If symmetry arguments are used to avoid repeats in the enumeration of the SSS’s for every ICS, the number of functional evaluations that are required within the exhaustive evaluation is

\[
M_{EE}(Q) = \sum_{l=2}^{Q} \sum_{i=1}^{N_{GCD}(l)} \sum_{n=1}^{N_{SSS}^{(l)}(N,x_n)} \tilde{T}_{SSS}^{(l)}(N,x_n)
\]

(17)

To compute \( M_{EE}(Q) \) we set the following approximate expression:

\[
M_{EE}(Q) \approx \sum_{l=2}^{Q} \sum_{i=1}^{N_{GCD}(l)} \tilde{T}_{SSS}^{(l)}(N,x_n)
\]

(18)

and for \( \tilde{T}_{SSS}^{(l)}(N) \), here, we make use of the exponential extrapolation of the number of SSS’s obtained examining the list of configurations that we directly enumerated up to 20 sites per cell [see Fig. 2(b)]. In Fig. 3 we display \( M_{EE} \) as a function of \( Q \); one notices that the exhaustive enumeration is already very demanding when \( Q = 20 \), a case for which \( N \approx 3 \times 10^6 \) structures need to be evaluated.

**B. Sampling of the configurational space for \( N \leq Q \)**

A sampling algorithm will explore the search space calculating a fraction of the total number of possible configurations. Such an approach will be advantageous if the number of evaluations needed to find the optimal configurations is much smaller than the total number of possible configurations. Thus far, sampling methods have been generally applied to a cell of fixed shape and fixed concentration containing \( N \) atoms, when the underlying crystal lattice is fixed (e.g., fcc). In the language of Fig. 1 this space contains all \( N_{SSS}(N,x_n) \) SSS’s corresponding to one ICS and to a given concentration \( x_n \).
IV. COMBINING DIRECT ENUMERATION OF ICS’S WITH SAMPLING OF SSS’S: THE “DIFFERENT SAMPLING FOR DIFFERENT CELL SHAPES” (DSDS) APPROACH

Instead of evaluating $P(\sigma)$ for all the SSS’s and all ICS’s, one could perform a sampling search for all SSS’s and repeat it systematically for each composition and each ICS (see Fig. 1). Such a strategy will provide for each given cell shape and composition the configuration with an extremal value of the functional $P$. From this standpoint, the combined approach, when all ICS’s are systematically considered, allows discovery of the configurations with extremal values of the functional $P(\sigma)$ that are in agreement with those found by a scan of the entire configurational space, i.e., based on an exhaustive evaluation. The advantage with respect to the exhaustive evaluation is clear: there is no need to enumerate and calculate all the SSS’s. We will refer to this combined approach as “Different Sampling for Different (cell) Shapes” (DSDS).

We next evaluate the computational cost, within the DSDS approach, of searching for the configuration $\sigma^*$ of extremal $P(\sigma)$ for each concentration $x_n$ represented in the space of all binary crystal compounds having a number of atoms per cell lower or equal to $Q$. We first evaluate the computational cost of determining $P(\sigma)$, using DSDS based on the GA, considering all SSS’s belonging to one given cell shape and all the related concentrations. The answer is $N_{\text{eval}}^{\text{GA}}(N,x_n) \times (N-1)$, where $N_{\text{eval}}^{\text{GA}}(N,x_n)$ is the number of evaluations of the physical functional $P(\sigma)$ needed by the GA to find the extremal configuration. $N_{\text{eval}}^{\text{GA}}(N,x_n)$ depends on the number $N$ of lattice sites (i.e., the length of the genome’s string$^{22}$) and on the concentration $x_n$. Furthermore, this search must be performed for each $N \leq Q$ and all the related ICS’s. Therefore, the number of evaluations of the physical functional $P(\sigma)$, that are needed to complete the search up to $Q$ atoms per cell using DSDS, is

$$M_{\text{DSDS}}(Q) = \sum_{i=2}^{Q} \sum_{l=1}^{N_{\text{ICS}}(l)} \left( \sum_{n=1}^{l-1} N_{\text{eval}}^{\text{GA}}(l,x_n) \right).$$

(19)

A lowest $N_{\text{eval}}^{\text{GA}}(N,x_n)$ could be achieved by the optimization of the parameters such as the number of genomes in the population, the mutation rate, and the survival rate that control the evolution of the genetic algorithm.$^{22}$ Usually, $N_{\text{eval}}^{\text{GA}}(N,x_n)$ is much smaller than the total number of configurations associated with a given ICS. However, when the GA search is performed on small combinatorial spaces, the algorithm easily ends up surveying equivalent configurations repeatedly and can accept them in the population as long as they meet the fitness criterion. This fact not only leads to a poor degree of diversity in the population but also, in the case of ICS’s containing $N \lesssim 10$ atoms, to a total number of evaluations of the physical functional that may be comparable to or even larger than the number of nonequivalent structures in the space explored. To evaluate $M_{\text{DSDS}}(Q)$ one needs to specify in Eq. (19) how $N_{\text{eval}}^{\text{GA}}(N,x_n)$ depends on $N$ and $x_n$. As this function is not known analytically, an approximation is needed for it. Here, we make the simple assumption that a maximum of 1000 functional evaluations are performed for each GA run independently of the number of atoms in the cell and of the concentration $x_n$. Then, under such assumption we have

$$M_{\text{DSDS}}^{\text{GA}}(Q) = 1000 \times \sum_{l=2}^{Q} N_{\text{ICS}}(l) \left( \sum_{n=1}^{l-1} 1 \right).$$

(20)

We notice that Eq. (20) becomes exact when one performs the search assigning a maximal number of functional evaluations, for instance 1000, to each GA search. Assuming $N_{\text{eval}}^{\text{GA}}(N,x_n)=1000$ in the approximate relation of Eq. (20) leads to overestimate the number of functional evaluations for small $Q$, typically for $Q < 16$. In Fig. 3, we compare the computational behavior of $M_{\text{DSDS}}^{\text{GA}}$ obtained using Eq. (20), and $M_{\text{EE}}$ as a function of $Q$. Interestingly, we observe that the DSDS search outperforms the exhaustive evaluation when $Q > 23$. The fact that the computational cost of the DSDS is lower than that of EE for $Q$ larger than a certain threshold value, is a result that does not depend on the assumption we made on $N_{\text{eval}}^{\text{GA}}(N,x_n)$. Indeed, this is due to the fact that the number of configurations that GA has to sample to find the extremum within a large combinatorial space is much smaller than the number of configurations in the space.

V. APPLICATION OF THE DSDS TO GROUND STATE SEARCH OF METALLIC ALLOYS

A. The cluster expansion

Given the atomic configuration $\sigma$ of an alloy defined as the set of $N$ lattice variables $\{S_1, S_2, \ldots, S_N\}$ the corresponding formation energy $\Delta H_{\text{CE}}(\sigma)$ can be written as an Ising-like cluster expansion$^{24}$ (CE)

$$\Delta H_{\text{CE}}(\sigma) = J_0 + \sum_i J_i S_i + \sum_{i<j} J_{ij} S_i S_j + \sum_{i<j<k} J_{ijk} S_i S_j S_k + \cdots.$$  

(21)

Equation (21) can be written$^{35}$ in terms of $n_{\text{pairs}}$ pairs and $N_{\text{MB}}$ many-body figures $f$

$$\Delta H_{\text{CE}}(\sigma) = E_{\text{CS}}(\sigma, x) + J_0 + (2x-1) J_1 + \sum_{\text{pairs}} n_{\text{pairs}} J_{\text{pair}} D_{\text{pair}} \Pi_{\text{pair}}(\sigma) + \sum_{\text{MB}} J_{\text{MB}} D_{\text{MB}} \Pi_{\text{MB}}(\sigma),$$

(22)

where $J_f$ are the interaction parameters and $D_f$ the symmetry degeneracy for each inequivalent figure $f$. The function $\Pi_f(\sigma)$ is the space-averaged product of spins at the vertices of the figure $f$ and therefore depends on the alloy configuration $\sigma$. The term $E_{\text{CS}}$ is the constituent energy (see Ref. 25). This term is included to remove the $k \rightarrow 0$ singularity of the strain energy. The interaction coefficients $J_f$ are obtained by mapping $\Delta H_{\text{CE}}(\sigma)$ onto an appropriate input base of fully relaxed $\Delta H_{\text{DFT}}(\sigma)$ energies, calculated, for instance within the local-density approximation to the density functional theory for a set of $N_\sigma$ structures. In principle, an infinite number of pair interactions and many-body figures should be used in the cluster expansion, whereas in practice the fast
convergence of the expansion with respect to the length of interaction makes a truncated expansion a satisfactory approximation for the energetics of the alloy. Recently, a procedure has been formulated to identify and calculate the pairs and the many-body interactions that are to be kept for the truncated cluster expansion to have real predictive power.\textsuperscript{17,18} Once a cluster expansion is established, one may want to identify the alloy configurations of relative stability. To accomplish this task the capability of surveying exhaustively and rapidly large sets of crystal structures is of fundamental importance. To identify ground states in principle one should perform a search that includes structures that contain a number of atoms per cell as large as possible. We saw that in practice a complete enumeration is affordable up to 20 atoms per cell. The DSDS method overcomes this practical limit allowing consideration of alloy structures with a number of atoms per cell larger than 20.

B. Ground state search of Au-Pd and Mo-Ta

Here, we perform a DSDS ground state search for the Au-Pd fcc alloy, for which a cluster expansion has been recently presented by Barabash et al. in Ref. 23, and for the Mo-Ta bcc alloy, using the cluster expansion obtained by Blum and Zunger in Ref. 26. The cluster expansion for Au-Pd includes thirteen pair-interaction and seven many-body figures of which five are three-body interactions and the other two four-body interactions. The cluster expansion for Mo-Ta employs eight pair interactions, four three-body figures and one four-body figure. The structures of minimum formation energy have been found by repeating the GA search at least four times starting from independent random initial populations. Following such a procedure, for every cell shape and concentration, one has a higher probability that the structure of minimum energy found by the repeated independent searches corresponds to the structure of lowest energy. From a cluster expansion standpoint, the maximal cell size above which it is reasonable to find out new ground states is dictated by the range of the interactions. Practically, however, we stop increasing the cell size when no new ground states are revealed, or if their depth is negligible.

1. Au-Pd

We start performing a DSDS ground state search restricted to the structures with up to 20 atoms per cell. We compare the ground state line with that obtained by exhaustive evaluation in Ref. 23: the ground state lines obtained by using the two different approaches are identical. Then, we proceed with the DSDS search on the set of crystal structures that have up to 32 lattice sites in the unit cell: the result of this search is shown in Fig. 4\textsuperscript{a}. The ground state line obtained by searching all the structures with up to 32 atoms per cell is compared in Fig. 4\textsuperscript{b} with that found searching structures with up to 20 atoms per cell. Interestingly, the ground state line obtained by the former search presents two breaking points at Pd concentration \( x = 1/6 \) and \( x = 3/8 \) that are new with respect to those already found searching up to 20 atoms (see Ref. 23). The structures of these two new ground states are shown in Fig. 5. The depth (see Ref. 16 for the

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig4.png}
\caption{(Color online) (a) The result of the DSDS ground state search for Au\textsubscript{1-x}Pd\textsubscript{x} performed up to 32 atoms per cell: each black circle represents the minimum energy found at the corresponding concentration \( x \) for a certain ICS. The red line is the ground state line and the empty red circles indicate the breaking points. (b) Comparison of the ground state energies of Au\textsubscript{1-x}Pd\textsubscript{x} obtained by exhaustive evaluation up to 20 atoms per cell (empty circles) and (filled circles) the DSDS search extended up to 32 atoms per cell. (c) Results of the ground state search for Mo\textsubscript{1-x}Ta\textsubscript{x} performed up to 32 atoms per cell using DSDS.}
\end{figure}
definition of this quantity) of these ground states relative to the tie line is, however, very small, i.e., 0.1 meV/atom, so that the ground state appears almost unchanged with respect to that obtained considering structures with no more than 20 atoms per cell. 0.1 meV/atom depth relative to the tie line of the neighboring ordered structures is indeed small, but within the numerical precision of the cluster expansion. Note that this is not the depth with respect to the random alloy which decides \( T_c \). At concentration \( x = 3/8 \) the structure obtained is not a superlattice, while at concentration \( x = 1/6 \) we obtain the superlattice \( (AB_3)(AB_5)(AB_3)(AB_5) \). Finding a superlattice along the \( (310) \) direction is in agreement with the picture that the exhaustive evaluation\(^{23} \) suggested for this region, which is characterized by the so-called “adaptive” structures. Such structures allow for efficient adaptation to infinitely small changes in the concentration within a fairly large range of alloy compositions. Such a property arises from the balance between the strain energy, which is favorably low along specific directions [that, in Au-Pd, are \( (100) \) and \( (310) \)], and the nearest-neighbor interaction energy, that is repulsive for same-type atoms: this makes unfavorable the formation of atomic planes composed of a single species and that include a large number of nearest-neighbor interactions. The lowest energy structure found by searching all structures with no more than 20 atoms per cell, for \( x = 1/6 \), is the superlattice \( (310) AB_5 \). Interestingly, the \( AB_5 \) structural motif is also present in the superlattice found by DSDS. Therefore, the superlattice configuration retrieved by DSDS can be seen as being derived from the \( AB_5 \) superlattice through a fluctuation in the number of \( B \) planes in the \( B \) rich region, so as to have along the \( (310) \) direction structural segments of type \( AB_4 \) or \( AB_6 \) that alternate with segments of type \( AB_5 \). The large unit cell structure that DSDS provides at composition \( x = 1/6 \) suggests the following hypothesis on the “adaptive” structures zone of the phase diagram: The superlattice \( (310)(AB_5)(AB_3)(AB_5)(AB_5) \) can be seen as an example of a more general class of superstructures that show a sequence of elemental motifs \( AB^n \), oriented along the \( (310) \) direction, with \( n \) not constrained to be fixed throughout the sequence [e.g., \( (AB_3)(AB_5)(AB_3) \)]. Such type of superstructures can accommodate infinitesimal changes in the concentration by varying the number of \( B \) planes of the pure \( B \) zone in the elemental motif \( AB^n \).

2. Mo-Ta

Mo-Ta represents a very interesting example of a bcc metallic alloy that shows a rich ground state line, as Blum and Zunger\(^ {18,26} \) pointed out by a ground state search based on exhaustive enumeration up to 20 atoms per cell. The DSDS search extended to structures containing up to 32 atoms per cell [see Fig. 4(c)] did not provide with new ground states with respect to the search limited to structures with not more than 20 atoms per unit cell. Therefore, the result of the DSDS search on this system shows the robustness of the small cell ground states predicted by exhaustive evaluation in Refs. 18 and 26.

The fact that new ground states with \( N > 20 \) have been found in Au-Pd and not in Mo-Ta may be explained considering the range of the interactions used in the cluster expansion that is longer for the Au-Pd cluster expansion than for the Mo-Ta cluster expansion.

VI. CONCLUSIONS

In this work we have presented a method for generating the configurational space of alloys and searching this space for the structures \( \sigma \) that show specific values of the physical functional \( P(\sigma) \). Our method takes advantage of a simple and fast procedure to exhaustively enumerate the ICS’s and combine this procedure with a robust sampling algorithm, such as the genetic algorithm, to obtain, among the SSS’s associated to each ICS, the configuration that presents the target value of \( P(\sigma) \). We used the DSDS method to search for the ground states of the Au-Pd and Mo-Ta alloy systems. In the Au-Pd case, we obtained ground states that were not predicted by previous exhaustive enumeration. In particular DSDS allowed discovery of a structural motif for the adaptive structures that was not anticipated by searching configurations with no more than 20 atoms per cell.

ACKNOWLEDGMENTS

We thank S. Barabash for help in using the cluster expansion of Eq. (21) for the Au-Pd system. This work was funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, under Contract No. DE-AC36-99GO10337 to NREL.


