



Cluster expansion and the configurational theory of alloys

J. M. Sanchez

Texas Materials Institute, The University of Texas at Austin, Austin, Texas 78712, USA

(Received 8 March 2010; published 4 June 2010)

A rigorous mathematical foundation for the cluster-expansion method is presented. It is shown that the cluster basis developed by Sanchez *et al.* [Physica A **128**, 334 (1984)] is a multidimensional discrete Fourier transform while the general formalism of Sanchez [Phys. Rev. B **48**, 14013 (1993)] corresponds to a multidimensional discrete wavelet transform. For functions that depend nonlinearly on the concentration, it is shown that the cluster basis corresponding to a multidimensional discrete Fourier transform does not converge, as it is usually assumed, to a finite cluster expansion or to an Ising-type model representation of the energy of formation of alloys. The multidimensional wavelet transform, based on a variable basis cluster expansion, is shown to provide a satisfactory solution to the deficiencies of the discrete Fourier-transform approach. Several examples aimed at illustrating the main findings and conclusions of this work are given.

DOI: [10.1103/PhysRevB.81.224202](https://doi.org/10.1103/PhysRevB.81.224202)

PACS number(s): 64.60.Cn, 71.15.Nc, 81.30.Bx

I. INTRODUCTION

An important goal of the theory of statistical thermodynamic is to provide a full and complete description of the state of compositional short- and long-range order in alloys.¹ Advances in both the theoretical and experimental fronts have depended heavily on the availability of a formal framework for the representation of functions of configurations such as the energy of alloy formation, observed x-ray diffuse scattering, band gaps in semiconductors, etc.

The use of an orthogonal and complete cluster basis to describe such functions was originally developed for the calculation of alloy phase diagrams,^{2–4} which were based on empirical Ising-type models for the energy and on the cluster-variation method⁵ for the treatment of the configurational entropy. The success of such calculations in reproducing general features of experimentally observed phase diagrams spurred interest in the development of first-principles approaches to alloy thermodynamics. In particular, Connolly and Williams⁶ proposed to combine the cluster-basis description with first-principles total-energy calculations of ordered compounds so as to obtain a representation of the energy of formation of alloys in any state of order. The promise to establish a bridge between *ab initio* alloy theory with statistical thermodynamic launched several efforts to formalize the approach.

A framework for the implementation of the method of Connolly and Williams⁶ was introduced by Sanchez *et al.*⁷ in 1984. The foundation of the formalism is the construction of a complete and orthonormal cluster basis in configurational space, which allows the description of functions of configurations in terms of cluster-expansion (CE) coefficients. The approach, known as the CE method, has been used extensively in combination with first-principles calculations of physical properties of metallic and semiconductor alloys. For the case of the energy of alloy formation, the method casts the energy in the form of an Ising-type model with constant expansion coefficients.

An important development in the evolution of the CE was the introduction by Asta *et al.*⁸ and Wolverton *et al.*⁹ of basis functions that resulted in concentration-dependent effective

cluster interactions. The results presented in Refs. 8 and 9 laid the ground for the development of a general formalism for the cluster expansion by Sanchez in 1993.¹⁰

However, the general theory of the CE (Ref. 10) was used by Gonis *et al.*¹¹ to raise serious questions regarding the validity of the expansion in the thermodynamic limit. The main conclusion of Ref. 11 is that energy of formation of alloys *cannot* be reduced to an Ising-type model with concentration-independent effective cluster interactions. Interestingly enough, the work embodied in Refs. 8–11 has not been addressed, one way or the other, by the community of practitioners of the CE which has consistently, and exclusively, continued to use the basis introduced by Sanchez *et al.*⁷ to expand a variety of physical quantities in order to infer, and even predict, complex alloy behavior. At this point in time, the conclusions arrived at by Gonis *et al.*¹¹ bring into question the validity of a significant body of work accumulated over more than 20 years, and solidly incorporated into the scientific literature.

Thus, one goal of this paper is to clarify a state of affairs concerning the CE method that, at best, can be characterized as confusing. To summarize our main results, we show that the CE is a rigorous mathematical formalism for the representation of functions of configurations. In particular, we show that an expansion in the basis developed in Ref. 7 corresponds, exactly, to a multidimensional Hadamard transform¹² and, therefore, to a multidimensional discrete Fourier transform (MDFT). In this basis, functions of configurations are represented by a cluster spectrum. Likewise, the general formalism of Sanchez¹⁰ corresponds to a multidimensional discrete wavelet transform that resolves functions of configuration in terms of their cluster spectrum and their concentration.

The fact that the cluster expansion is on a rigorous mathematical foundation implies, of course, that some of the results of Gonis *et al.*¹¹ are incorrect. The formalism developed here will show that the mathematical errors of Gonis *et al.*¹¹ are simply related to the fact that one must exercise some care in taking the limit of sums over an infinite number of clusters. Nevertheless, the main conclusion of Gonis *et al.*,¹¹ namely, that the energy of formation of alloys cannot be cast

in the form of an Ising-type model, is correct. The main culprit is, somewhat trivially, the nonlinear concentration dependence of the energy of formation of the *random* alloy which, in the thermodynamic limit, introduces an infinite number of vanishingly small Fourier-transform coefficients or, in the terminology of the CE, an infinite number of expansion coefficients in the basis of Ref. 7. The end result is that the energy of formation in the basis of Sanchez *et al.*⁷ does not reduce to a finite sum of cluster coefficients, i.e., to an Ising-type model. In this basis, the representation of the energy of formation as a CE, i.e., as a MDFT, is of course correct although of no obvious practical value.

We show that the general formalism of Sanchez,¹⁰ which leads to a multidimensional discrete wavelet transform (MDWT), provides a simple resolution to the shortcomings of the multidimensional discrete Fourier transform by resolving the energy into two components, the random energy and the ordering energy. The random energy depends only on concentration while the ordering energy depends on short- and long-range order. The ability of the MDWT to resolve the cluster spectrum and the concentration dependence of the ordering energy results in a CE with concentration-dependent effective cluster interactions. However, it is also possible to use a MDFT to expand the *ordering* energy in the basis of Ref. 7, in which case its representation is in the form of an Ising-type model.

The organization of the paper is as follows: we begin in Sec. II with a brief review of the general formalism of the CE as developed in Ref. 10 and include the formulation of the method in vector space in Sec. II B. Section III is devoted to the interpretation of the CE as a multidimensional discrete Fourier transform and as a multidimensional discrete wavelet transform. The method, which is applicable to any topology of the underlying lattice, is adapted to periodic crystals in Sec. IV. Brief comments connecting the CE with the general theory of statistical thermodynamics of alloys are presented in Sec. V. The implementation of the inversion method of Connolly and Williams⁶ is treated in Sec. VI. We conclude with examples in Sec. VII and a summary of the main results in Sec. VIII.

II. CLUSTER EXPANSION FORMALISM

In this section we briefly review the CE formalism developed in Ref. 10. We will restrict our discussion to binary systems although the basic CE framework can be extended to multicomponent alloys. Thus, we consider an N -point cluster consisting of two type of atoms, A and B. The type of atom at point p in the cluster is given by the spin variable σ_p which equals 1 or -1 for A and B atoms, respectively. A configuration of the system is given by $\vec{\sigma} = \{\sigma_1, \sigma_2, \dots, \sigma_N\}$. The fundamental idea behind the CE is to construct a set of basis functions for the 2×2 configurational space corresponding to $N=1$, and to build a set of basis functions for an arbitrary value of N by carrying out a direct product of the $N=1$ basis over all the points p in the cluster.

For $N=1$, the basis functions are $\{1, \varphi^\mu(\sigma)\}$ with¹⁰

$$\varphi^\mu(\sigma) = \frac{\sigma - \tanh(\mu)}{\sqrt{1 - \tanh(\mu)^2}}, \quad (1)$$

where $\sigma = \pm 1$ and μ is a real number. The meaning of μ and the important role it plays in the convergence of the CE will be discussed in Sec. III B.

The functions $\{1, \varphi^\mu(\sigma)\}$ are such that they form a complete and orthonormal basis under a scalar product defined by

$$\langle f, g \rangle_\mu = \sum_{\sigma=\pm 1} \left[\frac{e^{\mu\sigma}}{2 \cosh(\mu)} \right] f(\sigma)g(\sigma), \quad (2)$$

where $f(\sigma)$ and $g(\sigma)$ are functions in the two-dimensional space. The orthogonality and completeness of the basis are expressed by

$$\langle 1, \varphi^\mu \rangle_\mu = 0, \quad (3)$$

$$\langle 1, 1 \rangle_\mu = \langle \varphi^\mu, \varphi^\mu \rangle_\mu = 1, \quad (4)$$

$$1 + \varphi^\mu(\sigma)\varphi^\mu(\sigma') = \delta_{\sigma,\sigma'} \frac{2 \cosh(\mu)}{e^{\mu\sigma}}. \quad (5)$$

For a general N -point cluster the corresponding complete and orthonormal basis set in configurational space, of dimension 2^N , is obtained by carrying out the direct product of the two-dimensional basis functions $\{1, \varphi^\mu(\sigma)\}$. The basis functions in question are

$$\phi_\alpha^\mu(\vec{\sigma}_\alpha) = \prod_{p \in \alpha} \varphi^\mu(\sigma_p), \quad (6)$$

where the subindex α refers to one of the 2^N clusters of points in the N -point cluster and $\vec{\sigma}_\alpha$ is the set of spin variables contained in cluster α . The set includes the empty cluster ($\alpha=0$) for which $\phi_0^\mu(\vec{\sigma})=1$. A given cluster α will be labeled by a set of indices $\alpha = \{\eta, p, \nu\}$, where η denotes the type of cluster [e.g., point, nearest-neighbor (nn) pairs, etc.], p refers to its location, and ν to its orientation in the N -point cluster.

The scalar product of Eq. (2) becomes

$$\langle f, g \rangle_\mu = \sum_{\vec{\sigma}} \mathcal{K}_N^\mu[x(\vec{\sigma})] f(\vec{\sigma}) g(\vec{\sigma}) \quad (7)$$

with

$$\mathcal{K}_N^\mu(x) = \left[\frac{e^{\mu x}}{2 \cosh(\mu)} \right]^N, \quad (8)$$

and, where

$$x(\vec{\sigma}) = \frac{1}{N} \sum_{p=1}^N \sigma_p = \frac{N_A - N_B}{N}, \quad (9)$$

with N_A and N_B the number of A and B atoms in configuration $\vec{\sigma}$, respectively.

The orthonormality and completeness of the basis functions $\{\phi_\alpha^\mu(\vec{\sigma})\}$ follow from the corresponding properties of the basis $\{1, \varphi^\mu(\sigma)\}$ and are expressed by

$$\langle \phi_\alpha^\mu, \phi_\beta^\mu \rangle_\mu = \sum_{\vec{\sigma}} \mathcal{K}_N^\mu[x(\vec{\sigma})] \phi_\alpha^\mu(\vec{\sigma}_\alpha) \phi_\beta^\mu(\vec{\sigma}_\beta) = \delta_{\alpha,\beta} \quad (10)$$

and

$$\sum_\alpha \phi_\alpha^\mu(\vec{\sigma}_\alpha) \phi_\alpha^\mu(\vec{\sigma}'_\alpha) = \prod_{p=1}^N [1 + \varphi^\mu(\sigma_p) \varphi^\mu(\sigma'_p)] = \frac{\delta_{\vec{\sigma}, \vec{\sigma}'}}{\mathcal{K}_N^\mu[x(\vec{\sigma})]}. \quad (11)$$

It follows from the orthonormality and completeness of the basis $\{\phi_\alpha^\mu(\vec{\sigma}_\alpha)\}$ that a function of configuration $f(\vec{\sigma})$ can be written as

$$f(\vec{\sigma}) = \sum_\alpha j_\alpha(\mu) \phi_\alpha^\mu(\vec{\sigma}_\alpha) \quad (12)$$

with

$$j_\alpha(\mu) = \langle f, \phi_\alpha^\mu \rangle_\mu. \quad (13)$$

This representation of $f(\vec{\sigma})$ is commonly referred to as the cluster expansion. We re-emphasize at this point that the parameter μ can be any real number and it generates an infinite number of basis sets $\{\phi_\alpha^\mu(\vec{\sigma}_\alpha)\}$, all of which are orthonormal and complete in the configurational space $\vec{\sigma}$.

As pointed out in Ref. 10, the scalar product of Eq. (7) can be written as

$$\langle f, g \rangle_\mu = \sum_{N_A=0}^N \mathcal{D}_N^\mu[x(\vec{\sigma})] (f, g)_x \quad (14)$$

with

$$\mathcal{D}_N^\mu(x) = \mathcal{K}_N^\mu(x) \frac{N!}{N_A! N_B!} \quad (15)$$

and, where $(f, g)_x$ stands for

$$(f, g)_x = \frac{N_A! N_B!}{N!} \sum_{\vec{\sigma}}' f(\vec{\sigma}) g(\vec{\sigma}) \quad (16)$$

with the sum over $\vec{\sigma}$ restricted to configurations such that $N_A - N_B = Nx$.

In the limit $N \rightarrow \infty$, we have,¹⁰

$$\lim_{N \rightarrow \infty} \mathcal{D}_N^\mu(x) = \delta_{x, \tanh(\mu)}, \quad (17)$$

where δ is the Kronecker delta. Thus, it follows from Eq. (14) that

$$\lim_{N \rightarrow \infty} \langle f, g \rangle_\mu = \lim_{N \rightarrow \infty} (f, g)_{\tanh(\mu)} \quad (18)$$

and, in particular, that

$$\lim_{N \rightarrow \infty} \langle \phi_\alpha^\mu, \phi_\beta^\mu \rangle_\mu = \lim_{N \rightarrow \infty} (\phi_\alpha^\mu, \phi_\beta^\mu)_{\tanh(\mu)}. \quad (19)$$

A. Transformation between basis sets

The basis sets for different values of μ can be shown to be related by a linear transformation.¹⁰ Using Eq. (12) to expand the functions $\phi_\alpha^\mu(\vec{\sigma}_\alpha)$ in the basis $\{\phi_\alpha^{\mu'}(\vec{\sigma}_\alpha)\}$, we can write

$$\phi_\alpha^\mu(\vec{\sigma}_\alpha) = \sum_\beta \langle \phi_\alpha^\mu, \phi_\beta^{\mu'} \rangle_{\mu'} \phi_\beta^{\mu'}(\vec{\sigma}_\beta). \quad (20)$$

The projections $\langle \phi_\alpha^\mu, \phi_\beta^{\mu'} \rangle_{\mu'}$ can be easily calculated using the definition of the scalar product given by Eq. (7) and are given by

$$A_{\alpha\beta}^{\mu\mu'} = \langle \phi_\alpha^\mu, \phi_\beta^{\mu'} \rangle_{\mu'} = \begin{cases} 0 & \text{for } \beta \not\subset \alpha \\ \left[\frac{\cosh(\mu)}{\cosh(\mu')} \right]^{n_\alpha} & \text{for } \beta = \alpha \\ a_{\alpha\beta}^{\mu\mu'} & \text{for } \beta \subset \alpha, \end{cases}$$

where the matrix elements $a_{\alpha\beta}^{\mu\mu'}$ are

$$a_{\alpha\beta}^{\mu\mu'} = \frac{[\cosh(\mu)]^{n_\alpha}}{[\cosh(\mu')]^{n_\beta}} (x_{\mu'} - x_\mu)^{(n_\alpha - n_\beta)} \quad (21)$$

and, where n_α and n_β are the number of points in clusters α and β , respectively, and $x_\mu = \tanh(\mu)$.

The transformation matrix $A^{\mu\mu'}$ is nonsymmetric and, by properly ordering the clusters such that for any cluster α all its subclusters are lower in the hierarchy, $A^{\mu\mu'}$ can be written as a lower triangular matrix. Thus, we can write Eq. (20) as

$$\phi_\alpha^\mu(\vec{\sigma}_\alpha) = \sum_{\beta \subseteq \alpha} A_{\alpha\beta}^{\mu\mu'} \phi_\beta^{\mu'}(\vec{\sigma}_\alpha), \quad (22)$$

where the sum is over clusters β contained or equal to the cluster α . We note that the matrix of the linear transformation between the functions $\{\phi_\alpha^\mu(\vec{\sigma}_\alpha)\}$ and $\{\phi_\alpha^{\mu'}(\vec{\sigma}_\alpha)\}$ is independent of N and, therefore, holds valid for infinite systems. It can be easily shown that the inverse of the transformation $A^{\mu\mu'}$ is $A^{\mu'\mu}$,

$$\sum_\gamma A_{\alpha\gamma}^{\mu\mu'} A_{\gamma\beta}^{\mu'\mu} = \sum_{\alpha \supseteq \gamma \supseteq \beta} A_{\alpha\gamma}^{\mu\mu'} A_{\gamma\beta}^{\mu'\mu} = \delta_{\alpha,\beta}. \quad (23)$$

The linear relation between the basis sets for different values of μ implies that the expansion coefficients for different basis are also related by the same transformation,

$$\langle f, \phi_\alpha^\mu \rangle_\mu = \sum_{\beta \supseteq \alpha} A_{\beta\alpha}^{\mu'\mu} \langle f, \phi_\beta^{\mu'} \rangle_{\mu'}. \quad (24)$$

As we shall see in Sec. III B, the transformation $A^{\mu\mu'}$ plays an important role in the theory of the CE and, particularly, in the development of fast converging expansions.

B. Vector space representation

In this section we cast the CE formalism in terms of a vector space representation. The objective of the expansion is to describe functions of configurations of the form $f(\vec{\sigma}): \mathcal{V}_{\vec{\sigma}} \rightarrow \mathbb{R}$ between the 2^N configurational space $\mathcal{V}_{\vec{\sigma}}$ and the real numbers. Following standard procedures, we construct a linear transform that maps a 2^N input vector in the configurational space $\mathcal{V}_{\vec{\sigma}}$ into a linear combination of vectors in $\mathcal{V}_{\vec{\alpha}}$ also of dimension 2^N .

For $N=1$, the 2×2 transformation matrix is given by

$$\mathbf{M}_1^\mu = \frac{1}{\sqrt{2 \cos(\mu)}} \begin{pmatrix} e^{\mu/2} & 0 \\ 0 & e^{-\mu/2} \end{pmatrix} \begin{pmatrix} 1 & \varphi^\mu(1) \\ 1 & \varphi^\mu(-1) \end{pmatrix}. \quad (25)$$

The transform matrix \mathbf{M}_1^μ is such that

$$\mathbf{M}_1^\mu (\mathbf{M}_1^\mu)^T = (\mathbf{M}_1^\mu)^T \mathbf{M}_1^\mu = \mathbf{I}, \quad (26)$$

which follows from the orthogonality and completeness of the basis functions $\{1, \varphi(\sigma)\}$ given by Eqs. (3)–(5).

For a general value of N , the direct product of the 2×2 discrete spaces over all sites p results in a $2^N \times 2^N$ matrix representation \mathbf{M}_N^μ of the transform obeying the relations,

$$\mathbf{M}_N^\mu (\mathbf{M}_N^\mu)^T = (\mathbf{M}_N^\mu)^T \mathbf{M}_N^\mu = \mathbf{I} \quad (27)$$

with the rows of \mathbf{M}_N^μ given by vectors \vec{v}_σ^μ of the form

$$\vec{v}_\sigma^\mu = \sqrt{\mathcal{K}_N^\mu[x(\vec{\sigma})]} \vec{\phi}^\mu(\vec{\sigma}), \quad (28)$$

and where the components of $\vec{\phi}^\mu(\vec{\sigma})$ are the basis function $\{\phi_\alpha^\mu(\vec{\sigma}_\alpha)\}$ defined in Eq. (6). The matrix of the CE transform can be written as

$$\mathbf{M}_N^\mu = \Lambda_N^\mu \Phi_N^\mu, \quad (29)$$

where Λ_N^μ is a diagonal matrix with elements given by

$$(\Lambda_N^\mu)_{\vec{\sigma}, \vec{\sigma}} = \sqrt{\mathcal{K}_N^\mu[x(\vec{\sigma})]} \quad (30)$$

and where Φ_N^μ is a $2^N \times 2^N$ matrix with rows given by the vectors $\vec{\phi}^\mu(\vec{\sigma})$.

It follows from Eq. (27) that a function $f(\vec{\sigma})$, represented in the space $\mathcal{V}_{\vec{\sigma}}$ by a (row) vector \vec{f} , can be written as

$$\vec{f} = \vec{f} \mathbf{M}_N^\mu (\mathbf{M}_N^\mu)^T = (\vec{f} \Lambda_N^\mu \Phi_N^\mu \Lambda_N^\mu) (\Phi_N^\mu)^T. \quad (31)$$

The expansion coefficients defined in Eqs. (12) and (13) are given by the components of the vector $(\vec{f} \Lambda_N^\mu \Phi_N^\mu \Lambda_N^\mu)$, namely,

$$j_\alpha(\mu) = (\vec{f} \Lambda_N^\mu \Phi_N^\mu \Lambda_N^\mu)_\alpha. \quad (32)$$

Thus, the vector space representation of $f(\vec{\sigma})$ of Eq. (31) is identical to the expansion given by Eq. (12).

III. INTERPRETATION OF THE CLUSTER EXPANSION

The formalism presented in Sec. II provides a complete and rigorous formulation of the CE method previously developed in Ref. 10 and includes the basis of Sanchez *et al.*⁷ for $\mu=0$. In this section, we establish a direct connection between the CE for $\mu=0$ with the well-developed theory of the Hadamard transform which is itself equivalent to a multidimensional discrete Fourier transform of size $2^N=2 \times 2 \cdots \times 2$. We also show that by introducing the parameter μ , the CE expansion is closely related to a multidimensional wavelet transform.

A. Basis for $\mu=0$

The transformation corresponding to the cluster expansion for $\mu=0$, as well as the basis functions $\phi_\alpha^0(\vec{\sigma}_\alpha)$, are well known and widely used in applications such as signal pro-

cessing, data compression, and quantum algorithms. The transformation is known as the Hadamard¹² transform and the basis functions $\phi_\alpha^0(\vec{\sigma}_\alpha)$, which take values ± 1 , are Walsh functions of the discrete variable $\vec{\sigma}_\alpha$.¹²

Hadamard matrices \mathbf{H}_N of order 2^N can be constructed for every non-negative value of N . A well-known recursive algorithm to construct such matrices is

$$\mathbf{H}_N = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbf{H}_{N-1} & \mathbf{H}_{N-1} \\ \mathbf{H}_{N-1} & -\mathbf{H}_{N-1} \end{pmatrix}.$$

Thus, defining $\mathbf{H}_0=1$, the Hadamard matrix \mathbf{H}_1 is

$$\mathbf{H}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

It is also well known that \mathbf{H}_1 is equal to the matrix of the 2×2 discrete Fourier transform. Thus, a Hadamard transform of order N is exactly a MDFT of order $2^N=2 \times 2 \cdots \times 2$.

An alternative algorithm for the matrix element $(\mathbf{H}_N)_{i,j}$, with i and j running from 0 to 2^N-1 , uses the binary representation of the indices i and j ,

$$(\mathbf{H}_N)_{i,j} = \frac{1}{\sqrt{2^N}} (-1)^{\vec{i} \cdot \vec{j}}, \quad (33)$$

where $\vec{i} \cdot \vec{j}$ is the dot product of the binary representation of the numbers i and j [for example, for $i=4$ and $j=5$, $\vec{i} \cdot \vec{j} = (1, 0, 0) \cdot (1, 0, 1) = 1$].

Equation (33) for the i, j element of the Hadamard matrix follows from a one-to-one correspondence between the binary representation of the set of integers $i = \{0, 1, \dots, 2^N-1\}$ and the 2^N configurations $\vec{\sigma}$. For a given configuration $\vec{\sigma} = \{\sigma_1, \sigma_2, \dots, \sigma_N\}$, the string $\vec{i} = \{(1+\sigma_1)/2, (1+\sigma_2)/2, \dots, (1+\sigma_N)/2\}$ corresponds to a vector of length N with components equal to 0 or 1 and, therefore, to the binary representation of an integer i with $0 \leq i \leq 2^N-1$. Likewise, a correspondence can be established with the 2^N clusters α by assigning the set of positions $\{p\}$ of the nonzero elements in \vec{i} to a cluster α .

The i th row of the Hadamard matrix is then

$$\vec{h}_i = \frac{1}{\sqrt{2^N}} \{1, (-1)^{\vec{i} \cdot \vec{1}}, (-1)^{\vec{i} \cdot \vec{2}}, \dots, (-1)^{\vec{i} \cdot \vec{m}}\}, \quad (34)$$

where \vec{m} stands for the binary representation of 2^N-1 . Each component j in \vec{h}_i corresponds to the value of the basis functions $\phi_\alpha^0(\vec{\sigma}_\alpha)$ for cluster j in configuration i .

It follows from Eq. (26) that the matrix of the CE for $N=1$ and $\mu=0$, i.e., \mathbf{M}_1^0 , is identical to \mathbf{H}_1 . It is also straightforward to show that the transformation matrix corresponding to the CE for $\mu=0$ and any value of N , i.e., \mathbf{M}_N^0 , is equal to the matrix of the Hadamard transform of order N . Thus, the CE for $\mu=0$ is *exactly* a multidimensional discrete Fourier transform.

We expand briefly on the multidimensional DFT since it is a familiar and widely accepted approach. Consider a N -dimensional function $f(\vec{x}) = f(x_1, x_2, \dots, x_N)$ of the discrete

variables x_n . For each one of the N dimensions, the discrete variables take values $x_n = \{0, 1, \dots, M_n\}$. The multidimensional DFT is defined as

$$F(\vec{k}) = \frac{1}{\prod_{j=1}^N M_j} \sum_{n_1=0}^{M_1-1} \dots \sum_{n_N=0}^{M_N-1} e^{-2\pi i \vec{k} \cdot (\vec{x}/\vec{M})} f(\vec{x}), \quad (35)$$

where \vec{x}/\vec{M} stands for

$$\vec{x}/\vec{M} = \{x_1/M_1, x_2/M_2, \dots, x_N/M_N\},$$

and the inverse transformation is

$$f(\vec{x}) = \sum_{k_1=0}^{M_1-1} \dots \sum_{k_N=0}^{M_N-1} e^{2\pi i \vec{x} \cdot (\vec{k}/\vec{M})} F(\vec{k}). \quad (36)$$

We can translate the multidimensional DFT into the notation of the cluster expansion as follows. The number of different dimensions of the total space is the number of points p in the system, i.e., N , and the discrete variables x_p are $x_p = (1 + \sigma_p)/2$ with $p = 1, 2, \dots, N$. The x_p take values 0 or 1 and, therefore, $M_n = 2$ for all subspaces $n = 1, 2, \dots, N$ of the 2^N -dimensional space $\vec{\sigma}$. Thus, the DFT of Eq. (35) becomes

$$F(\vec{k}) = \frac{1}{2^N} \sum_{\vec{\sigma}} (-1)^{\vec{k} \cdot \vec{x}} f(\vec{\sigma}) \quad (37)$$

and the inverse DFT,

$$f(\vec{x}) = \sum_{\vec{k}} (-1)^{\vec{k} \cdot \vec{x}} F(\vec{k}). \quad (38)$$

The vector $\vec{k} = \{k_j\}$, with N components k_j that take values 0 or 1, corresponds to a cluster α given by the set of positions $\{p\}$ in \vec{k} with $k_j \neq 0$. In Eqs. (37) and (38), $(-1)^{\vec{k} \cdot \vec{x}}$ is the value of value of the basis function $\phi_\alpha^0(\vec{\sigma}_\alpha)$ for cluster \vec{k} , i.e. α , in configuration \vec{x} , i.e. $\vec{\sigma}$. Therefore, Eq. (37) is

$$F(\vec{k}) \equiv \frac{1}{2^N} \sum_{\vec{\sigma}} f(\vec{\sigma}) \phi_\alpha^0(\vec{\sigma}_\alpha) = \langle f, \phi_\alpha^0 \rangle_0 = j_\alpha(\mu) \quad (39)$$

and the inverse DFT of Eq. (38),

$$f(\vec{x}) \equiv f(\vec{\sigma}) = \sum_{\alpha} \langle f, \phi_\alpha^0 \rangle_0 \phi_\alpha^0(\vec{\sigma}_\alpha), \quad (40)$$

which is identical to the expansion of Eq. (12) for $\mu = 0$.

In the limit of $N \rightarrow \infty$, the multidimensional DFT becomes a continuous transformation described by a Fourier-transform operator. However, for practical applications, the standard procedure is to work with the discrete version of the Fourier transform and increase the number of waves, or clusters, so as to obtain a good representation of the function of interest. It is not uncommon to encounter functions that require a very large number of Fourier components, possibly infinite, in order to achieve an accurate representation of the function. In such cases, although the Fourier transform is of course valid, it does not provide a particularly useful representation. This is precisely the case for the energy of formation of binary alloys.

Thus, we next address the issue of whether or not it is possible to describe the energy of formation of a binary alloy using an Ising-type representation of the energy with a finite number of concentration-independent pair and multisite effective interactions. To that end, we consider a periodic lattice with N points and the function $f^{(2)}(\vec{\sigma}) = x^2(\vec{\sigma})$, where $x(\vec{\sigma})$ is the concentration of configuration $\vec{\sigma}$ given by Eq. (9). It follows trivially from Eq. (9) that the CE of $f^{(2)}(\vec{\sigma})$ is given by

$$f^{(2)}(\vec{\sigma}) = \frac{1}{N} + \frac{1}{N^2} \sum_p \sum_{p' \neq p} \sigma_p \sigma_{p'}, \quad (41)$$

where the sum is over all $N(N-1)$ pairs $\{p, p'\}$. The corresponding expansion coefficients are $j_0(0) = 1/N$ and $j_{\{p, p'\}}(0) = 1/N^2$.

It is straightforward to show [see Eq. (37)] that the multidimensional DFT of $f^{(2)}(\vec{\sigma})$ is given by

$$X(\vec{k}) = \frac{1}{N} \delta_{|\vec{k}|, 0} + \frac{1}{N^2} \delta_{|\vec{k}|, 2} \quad (42)$$

and, therefore,

$$f^{(2)}(\vec{\sigma}) = f^{(2)}(\vec{x}) = \frac{1}{N} + \frac{1}{N^2} \sum_{\vec{k}} \delta_{|\vec{k}|, 2} (-1)^{\vec{k} \cdot \vec{x}}. \quad (43)$$

Equations (41) and (43) are of course the same and show that the CE (or DFT) of $x^2(\vec{\sigma})$ involves an infinite sum over all $N(N-1)$ pairs (or vectors \vec{k} with $|\vec{k}|=2$). While the CE clearly converges in the limit $N \rightarrow \infty$, the sums in Eqs. (41) and (43) are on the order of N^2 and, therefore, functions that have a nonlinear dependence on $x(\vec{\sigma})$ cannot be described by a finite sum of Fourier transform or cluster coefficients. With regard to the energy of formation of binary alloys, such a nonlinear dependence in concentration is always present in the energy of the totally disordered or random configuration. Thus, an important and dominant component of the energy of formation, namely, the energy of the random alloy *cannot* be expressed in the form of an Ising-type model with concentration-independent interactions. This fact brings into question the validity of essentially all applications of the cluster expansion to date which, as mentioned in Sec. I, have exclusively used the CE in the basis with $\mu = 0$.

In the context of the CE, resolving the energy of formation of the random alloy from the rest of the energy, which is usually referred to as the ordering energy, is achieved in a straightforward manner by using a set of basis functions for the CE with a value of μ that is matched to the concentration of the alloy under consideration. This approach, discussed further in Sec. III B, leads to a representation of the energy of formation in terms of the energy of the random alloy, which depends only on concentration, plus the ordering energy which one might expect can be described by a rapidly convergent CE. The issue of whether the ordering energy can be represented by an Ising-type model, i.e., with concentration-independent expansion coefficients, will be addressed with the specific example of Fe-Co alloys in Sec. VII.

B. Basis $\mu \neq 0$

As shown in Sec. III A, the CE for $\mu=0$ is a multidimensional DFT. For $\mu \neq 0$, the CE correspond to the closely related MDWT of size $2 \times 2 \cdots \times 2$. Discrete wavelet transforms are commonly used in the analysis of time-dependent signals $f(t)$. While the Fourier transform allows for a full frequency resolution of $f(t)$, it provides no resolution in time and, therefore, no information as to what frequencies might be relevant at a given time t . The idea behind wavelet analysis is to use a scalable and modulated window that can be shifted along the time axis for different scales or sizes of the window and to determine the frequency spectrum as a function of the window's scale and translation. For each translation and scale, the result is a time-frequency representation of the signal $f(t)$.

In the terminology of binary alloys, the CE for $\mu=0$ allows perfect resolution of functions of configuration in terms of clusters (or frequencies) with the cluster spectrum given by the expansion coefficients $j_\alpha(0)$. Resolution in both clusters and concentration can be achieved by expressing a given function $f(\vec{\sigma})$ in the basis $\{\phi_\alpha^\mu(\vec{\sigma})\}$ with $\tanh(\mu)$ equal to the concentration x of configuration $\vec{\sigma}$. The functions $\phi_\alpha^\mu(\vec{\sigma})$ are then multidimensional wavelets, derived from the mother wavelets $\phi_\alpha^0(\vec{\sigma})$ by means of a translation of $\tanh(\mu)$ and a compression of $\cosh(\mu)$ in each of the 2×2 configurational subspaces.

It is instructive to see the effect of the MDWT on the expansion of the function $f^{(2)}=x^2(\vec{\sigma})$ considered in Sec. III A. For an arbitrary value of μ , the expansion coefficients $j_\alpha(\mu)$ can be easily obtained from the definition of the scalar product of Eq. (7). All expansion coefficients $j_\alpha(\mu)$ vanish identically except for the empty, point, and pair clusters. For the relevant clusters, the projections of $f^{(2)}(\vec{\sigma})$ are

$$j_0(\mu) = \langle f^{(2)}, 1 \rangle_\mu = x_\mu^2 + \frac{1-x_\mu^2}{N},$$

$$j_{\{p\}}(\mu) = \langle f^{(2)}, \phi_{\{p\}}^\mu \rangle_\mu = x_\mu \sqrt{1-x_\mu^2} \left(\frac{N-1}{N^2} \right),$$

$$j_{\{p,p'\}}(\mu) = \langle f^{(2)}, \phi_{\{p,p'\}}^\mu \rangle_\mu = \frac{1-x_\mu^2}{N^2},$$

where $x_\mu = \tanh(\mu)$ and $\{p\}$ and $\{p,p'\}$ refer to the point and pair clusters, respectively. The expansion coefficients for these clusters are independent of the location of the cluster in the lattice.

Thus, the full CE of $f^{(2)}(\vec{\sigma})$ in a basis with a fixed value of μ is

$$f^{(2)}(\vec{\sigma}) = x_\mu^2 + \frac{1-x_\mu^2}{N} + x_\mu \sqrt{1-x_\mu^2} \left(\frac{N-1}{N^2} \right) \sum_p \frac{(\sigma_p - x_\mu)}{\sqrt{1-x_\mu^2}} + \frac{1-x_\mu^2}{N^2} \sum_p \sum_{p' \neq p} \frac{(\sigma_p - x_\mu)(\sigma_{p'} - x_\mu)}{(1-x_\mu^2)}. \quad (44)$$

We note that for $x_\mu=0$ (i.e., $\mu=0$) we recover the expansion or DFT calculated in Sec. III A. We also note that for a fixed and arbitrary value of μ the sums over clusters in Eq.

(44) will in general diverge and the CE will not reduce to a sum over a finite number of terms. However, by matching the basis μ with the concentration $x(\vec{\sigma})$, i.e., making $x_\mu = x(\vec{\sigma})$, the CE (or MDWT) is

$$f^{(2)}(\vec{\sigma}) = x^2 + \frac{(1-x^2)}{N} \left[1 + \sum_n \omega_n \frac{(z_n - x^2)}{(1-x^2)} \right], \quad (45)$$

where z_n is the pair-correlation function for pair n and ω_n is the corresponding number of pairs per lattice point (i.e., half the coordination number of pair n). The quantity in the square brackets vanishes and, therefore, the MDWT reproduces $f^{(2)}$ exactly.

The results for the simple example of $f^{(2)}(\vec{\sigma})$ can be generalized by means of the transformation $A^{\mu\mu'}$ developed in Sec. II A. In particular, the CE coefficients $j_\alpha(\mu)$ in the basis set $\phi_\alpha^\mu(\vec{\sigma})$ are directly related to the coefficients $j_\alpha(0)$ for $\mu=0$ by the transformation $A^{\mu,0}$ [see Eq. (21)]. Thus, from Eq. (24) we obtain

$$j_\alpha(x) = (1-x^2)^{n_\alpha/2} \left[j_\alpha(0) + \sum_{\beta \supset \alpha} x^{(n_\beta - n_\alpha)} j_\beta(0) \right], \quad (46)$$

where n_α and n_β are the number of points in clusters α and β , respectively, and where we have made explicit the concentration dependence of $j_\alpha(x)$ by taking $x = \tanh(\mu)$.

Equation (46) is central to the CE. It expresses the CE coefficients at concentration x , $j_\alpha(x)$, as an infinite sum of the DFT or CE coefficients $j_\beta(0)$ for $\mu=0$. For a given cluster α with n_α points, the corresponding coefficient $j_\alpha(x)$ is given by a power series in x in which the coefficient of x^m is a sum over all clusters β of the bare DFT coefficients $j_\beta(0)$ with $n_\alpha + m$ points. Thus, by including higher values of m we can introduce into the CE the contributions to $j_\alpha(x)$ from all clusters that contain the cluster α .

We defer further discussion of Eq. (46) until we adapt the CE to the special case of periodic crystalline binary alloys. In this regard, we note that the formalism as developed so far applies to any topology of the N -point cluster and, thus, it is more general than needed for the usual application of the CE to crystalline systems. For the case of infinite crystalline lattices, or alternatively, systems with periodic boundary conditions, the space-group symmetry of the lattice results in a significantly simplified version of the cluster expansion that, in turn, can be interpreted in terms of a generalized statistical theory of alloys.

IV. PERIODIC LATTICES

Of particular interest is the application of the CE to functions $f(\vec{\sigma})$ for binary alloys defined on a periodic lattice with a given space-group symmetry. The symmetry operations of the space group of the undecorated lattice are of the form $(\theta|\tau)$, where θ is a point-group symmetry operation and τ is a lattice translation. For such systems, we label a cluster α by the indices $\{\eta, p, \nu_\eta\}$, where η refers to the type of cluster (e.g., shape and size), $p = \{1, 2, \dots, N\}$ is the location of the cluster in the lattice (e.g., the origin of the primitive unit cell containing the center of mass of the cluster), and $\nu_\eta = \{1, 2, \dots, \omega_\eta\}$ refers to one of the ω_η clusters in the orbit of

$\alpha=\{\eta, p, \nu_\eta\}$ generated by the point-group symmetry of the lattice. Thus, $N\omega_\eta$ is the total number of clusters of type η in the lattice.

Consider a space-group symmetry operation $(\theta|\tau)$ of the undecorated lattice acting on the crystal in a given configuration $\vec{\sigma}$. The configuration $\vec{\sigma}$ is transformed into a configuration $\vec{\sigma}'$ that is a permutation of the elements in $\vec{\sigma}$, a given cluster $\alpha=\{\eta, p, \nu_\eta\}$ transforms into $\alpha'=\{\eta, p', \nu'_\eta\}$, the basis functions $\phi_\alpha^\mu(\vec{\sigma}_\alpha)$ into $\phi_{\alpha'}^\mu(\vec{\sigma}'_{\alpha'})$ and $f(\vec{\sigma})$ into $f(\vec{\sigma}')$. Since the projection $\langle f, \phi_\alpha^\mu \rangle_\mu$ involves a sum over all configurations $\vec{\sigma}$ (or $\vec{\sigma}'$), it follows that in the absence of external fields, i.e., in the absence of spatial biases, the expansion coefficients of $f(\vec{\sigma})$ have the space-group symmetry of the underlying lattice and, therefore, are independent of the location p or orbit ν_η of cluster α ,

$$\langle f, \phi_\alpha^\mu \rangle_\mu = (\theta|\tau)\langle f, \phi_\alpha^\mu \rangle_\mu = \langle f, \phi_{\alpha'}^\mu \rangle_\mu. \quad (47)$$

Using the symmetry of the expansion coefficients, the cluster expansion of Eq. (12) can be written as

$$f(\vec{\sigma}) = J_0(\mu) + \sum_{\eta \neq 0} \omega_\eta J_\eta(\mu) z_\eta^\mu(\vec{\sigma}), \quad (48)$$

where

$$z_\eta^\mu(\vec{\sigma}) = \frac{1}{N\omega_\eta} \sum_{p=1}^N \sum_{\nu_\eta=1}^{\omega_\eta} \phi_{\alpha_\eta}^\mu(\vec{\sigma}_{\alpha_\eta}) \quad (49)$$

with the expansion coefficients given by

$$J_0(\mu) = \langle f, \phi_0^\mu \rangle_\mu = j_0(\mu) \quad (50)$$

and

$$J_\eta(\mu) = N \langle f, \phi_{\alpha_\eta}^\mu \rangle_\mu = N j_{\alpha_\eta}(\mu) \quad (51)$$

with $\alpha_\eta=\{\eta, p, \nu_\eta\}$.

A few comments concerning the functions $z_\eta^\mu(\vec{\sigma})$ and the coefficients $J_\eta(\mu)$ are in order. First we note that in the limit of $N \rightarrow \infty$ and for μ such that $\tanh(\mu)=x$, with x the concentration of configuration $\vec{\sigma}$, the $z_\eta^\mu(\vec{\sigma})$ become continuous variables that vanish for the random state and are such that $|z_\eta^\mu(\vec{\sigma})| \leq 1$. For the random state we specifically mean configurations for which the correlation functions in the basis $\mu=0$ are given by $z_\eta^0(\vec{\sigma})=x^{n_\eta}$ with n_η the number of points in cluster η . Furthermore, the variables $\{z_\eta^\mu\}$ are the specific versions of extensive variables $X_\eta^\mu=Nz_\eta^\mu$ with X_η^μ given by the unnormalized sum in Eq. (49).

With regard to the expansion coefficients, we note that the $j_\eta(\mu)$ for $\eta>0$ can be obtained by projecting the function $f(\vec{\sigma})$ onto any cluster α of type η irrespective of its location p or orientation ν_η in the lattice. We also see from Eq. (51) that the relevant CE coefficients for $\eta>0$ are $J_\eta(\mu) = N j_{\alpha_\eta}(\mu)$, showing that the projections $j_\alpha(\mu)$ are, at least, of order $1/N$.

From the linear relation between the $\phi_\alpha^\mu(\vec{\sigma}_\alpha)$ for different values of μ [see Eq. (22)], it follows that one set of variables $\{z_\eta^\mu\}$ can be written as a linear combination of any other set of variables $\{z_{\eta'}^\mu\}$,

$$z_\eta^\mu(\vec{\sigma}) = \sum_{\eta' \subseteq \eta} B_{\eta\eta'}^{\mu\mu'} z_{\eta'}^{\mu'}(\vec{\sigma}), \quad (52)$$

where the sum is over all subclusters of η , including $\eta'=\eta$ and $\eta'=0$. If the cluster of type η' is not a subcluster of η the coefficients $B_{\eta\eta'}^{\mu\mu'}$ equal 0, otherwise they are given by

$$B_{\eta\eta'}^{\mu\mu'} = N_{\eta'}^\eta \left(\frac{\omega_{\eta'}}{\omega_\eta} \right) A_{\eta\eta'}^{\mu\mu'} \quad (53)$$

with $N_{\eta'}^\eta$ the number of clusters of type η' in a cluster of type η and $A_{\eta\eta'}^{\mu\mu'}$ are defined in Eq. (21). It also follows from Eqs. (21), (52), and (53) that the inverse of $B^{\mu\mu'}$ is $B^{\mu'\mu}$.

As discussed in Sec. III B, one can achieve significant improvement in the convergence of the coefficients of the CE by matching the basis μ to the concentration x of the function under study. We will refer to this approach as the variable basis cluster expansion or VBCE. The general form of the expansion coefficients in the VBCE for periodic systems, i.e., the equivalent of Eq. (46), follows directly by applying the transformation $B^{\mu 0}$ to the DFT coefficients $J_\alpha(0)$. The result is

$$J_\eta(x) = (1-x^2)^{n_\eta/2} \sum_{m=0}^{m_\eta^{(0)}} x^m K_\eta^{(m)}, \quad (54)$$

where, as before, n_η is the number of points in the cluster of type η , x is the concentration of the configuration $\vec{\sigma}$ under consideration, $m_\eta^{(0)}$ is an integer that sets the maximum size $n_\eta+m_\eta^{(0)}$ of the clusters contributing to $J_\eta(x)$, and where $K_\eta^{(m)}$ is given by

$$K_\eta^{(m)} = \sum_{\eta'} N_{\eta'}^{\eta'} \left(\frac{\omega_{\eta'}}{\omega_\eta} \right) J_{\eta'}(0). \quad (55)$$

In Eq. (55) the sum over η' is over all clusters with $n_\eta+m$ points, $N_{\eta'}^{\eta'}$ is the number of η clusters contained in η' , and $J_{\eta'}(0)$ are the coefficients of the DFT or, equivalently, the CE coefficients for $\mu=0$. As we shall see in Sec. VII the coefficients $K_\eta^{(m)}$ are the parameters to be determined in a standard fitting of the CE coefficients to the energies of ordered compounds, which are typically calculated using *ab initio* methods.

V. CE AND THE STATISTICAL THEORY OF ALLOYS

The formalism developed for the CE identifies a complete set of configurational variables for any function of configuration. For systems defined on a periodic lattice, these configurational variables, which one may call canonical, are the set $\{z_\eta^\mu(\vec{\sigma})\}$ with $\tanh(\mu)=x(\vec{\sigma})$. As mentioned in the previous section, the variables in question are the specific versions of extensive variables $X_\eta=Nz_\eta^\mu$, they vanish for the random state and are such that $|z_\eta^\mu| \leq 1$. It also follows from the formalism that for the choice $\tanh(\mu)=x(\vec{\sigma})$, the variable associated with point cluster vanishes, i.e., $z_1^\mu=0$.

The CE is particularly well adapted to describe extensive functions in configurational space, such as the energy of for-

mation of alloys. With the choice of X_η as the configurational variables, an extensive function $E(\{X_\eta\})$ takes the form

$$E(\{X_\eta\}) = E_0(x) + \Delta E(\{X_\eta\}), \quad (56)$$

where $E_0(x)$ is the value of the function in the random state, which depends only on concentration, and where $\Delta E(\{X_\eta\})$ contains all information on the dependence of the function on short- and long-range order at a fixed concentration x . Both functions are extensive and, in particular, $\Delta E(\{X_\eta\})$, being a homogeneous function of degree 1, can be written as

$$\Delta E(\{X_\eta\}) = \sum_\eta \left(\frac{\partial \Delta E}{\partial X_\eta} \right)_{\vec{x}} X_\eta. \quad (57)$$

Therefore, the expansion coefficients $J_\eta(x)$ are the partial derivatives of $\Delta E(\{X_\eta\})$ with respect to X_η . It also follows from Euler's homogeneous functions theorem that

$$\vec{x} \cdot \nabla \left(\frac{\partial \Delta E}{\partial X_\eta} \right) = \sum_{\eta'} \left(\frac{\partial J_{\eta'}(x)}{\partial X_{\eta'}} \right)_{\vec{x}} X_{\eta'} = 0, \quad (58)$$

which holds for all expansion coefficients J_η . For a CE at a fixed value of μ , including $\mu=0$, Eq. (58) is trivially obeyed since the $J_\eta(\mu)$ are constants.

In the VBCE representation, Eq. (58) become,

$$z_1^\mu \left[\frac{\partial J_\eta(x)}{\partial z_1^\mu} \right]_{\vec{x}} = 0, \quad (59)$$

which are also trivially obeyed since $z_1^\mu \equiv 0$. However, Eq. (59) also shows that, for homogeneous functions of degree 1, the expansion coefficients $J_\eta(x)$ in the VBCE can, *in general*, be any function of concentration. We also note that in the VBCE, a Taylor expansion of $\Delta E(\{X_\eta\})$ about the random state, for which $X_\eta=0$, coincides with Euler's Eq. (57). Therefore, the expansion coefficients $J_\eta(x)$ are the partial derivatives of $\Delta E(\{X_\eta\})$ evaluated in the random state.

Thus, in summary, the VBCE provides a standard representation of extensive functions with a specific (and convenient) choice of configurational variables that vanish for the random state. The expansion coefficients $J_\eta(x)$ are at most function of concentration and correspond to the partial derivatives of $\Delta E(\{X_\eta\})$ with respect to X_η evaluated in the random state.

As shown in Sec. III A, a nonlinear dependence in concentration cannot be described by a finite set of coefficients using a DFT or, for that matter, the CE for any fixed value of μ . The problem is satisfactorily resolved using the MDWT or VBCE, and the solution is generally applicable to any extensive function. However, for intensive functions, one cannot rule out nonlinear dependences on the configurational variables z_η^μ . In such cases, both the regular CE and the VBCE require an infinite (or very large) number of expansion coefficients for a reasonably accurate representation of the function.

It should be apparent that the use of the CE (or the VBCE) to describe intensive functions needs to be approached carefully. In statistical thermodynamics, intensive functions of configuration are given by derivatives of extensive functions with respect to some appropriate intensive

variable. Thus, the obvious solution is to approach the expansion of intensive functions by expanding the relevant extensive function for different values of the relevant intensive variables and calculate the derivatives of the expansion coefficients themselves.

VI. INVERSION OF THE CE

A common application of the CE is to use it in conjunction with *ab initio* total-energy calculations to obtain the expansion coefficients or the effective cluster interactions. As mentioned in Sec. I, the method was first proposed by Connolly and Williams⁶ in 1983 and, since then, has been extensively used for metallic and semiconductor alloys. The essence of the approach is to calculate a set of energies $\vec{E} = \{E_n\}$ for n ordered structures and to fit the coefficients of the CE for a predetermined set of m clusters η . Irrespective of the basis functions used, the problem can always be cast in the form of a linear relation between the energies \vec{E} , a matrix \mathbf{Z} that characterizes the configurations of the n structures in the basis of choice, and a vector \vec{V} that contains all relevant information of the expansion coefficients for the selected clusters. Thus,

$$\vec{E} = \mathbf{Z}\vec{V}. \quad (60)$$

If, as it is generally practiced, the expansion is carried out in the basis $\mu=0$, then the dimension of the matrix \mathbf{Z} is $n \times m$. The rows of \mathbf{Z} are the vectors $\vec{z}_n = \{z_\eta^0(\vec{\sigma}_n)\}$ with $z_\eta^0(\vec{\sigma}_n)$ given by Eq. (49) and $\vec{\sigma}_n$ is the configuration of structure n with energy E_n . The components of the vector \vec{V} are then the CE coefficients $\omega_\eta J_\eta(0)$. The structure of Eq. (60) remains the same if one chooses a fixed value of $\mu \neq 0$.

In the VBCE described in Sec. IV, we use a different basis μ_n for each concentration x_n such that $\tanh(\mu_n) = x_n$. The components of the vector \vec{V} are the coefficients $K_\eta^{(m)}$ defined by Eq. (55) which are also linearly related to the energies \vec{E} by a matrix \mathbf{Z} . The vector \vec{V} is of length $m^{(0)}$ given by

$$m^{(0)} = \sum_{\eta=0}^{m-1} m_\eta^{(0)}. \quad (61)$$

It follows that the dimension of the matrix \mathbf{Z} is $n \times m^{(0)}$ with rows given by

$$\vec{z}_n = \{(1 - x_n^2)^{n/2} x_n^m z_\eta^{\mu_n}(\vec{\sigma}_n)\}, \quad (62)$$

where m_η runs from 0 to $m_\eta^{(0)}$ and η from 0 to $m-1$.

In most cases, the linear relation of Eq. (60) can only be solved approximately by means of a least-square fitting of the expansion coefficients \vec{V} to the input energies \vec{E} . There are, however, only two cases to consider and they are distinguished by the dimension d_\emptyset of the null space of the matrix \mathbf{Z} . If $d_\emptyset=0$, the solution to the least-square fitting is unique and is given by

$$\vec{V} = \mathbf{Z}^+ \vec{E}, \quad (63)$$

where \mathbf{Z}^+ is the pseudoinverse of the matrix \mathbf{Z} .

For the case $d_\emptyset > 0$, the linear relation in Eq. (60) is overdetermined and the least-square fitting problem has an infi-

nite number of solutions. Thus, in such cases, it is customary to introduce some *ad hoc* assumption regarding the behavior of the expansion coefficients in order to select one solution, say \vec{V}_0 , out of an infinite number of them. For example, for such an overdetermined linear problem, Eq. (63) gives the least-square solution for the case in which the norm of \vec{V} is minimum. Irrespective of the conditions imposed on the least-square fitting algorithm, they all lead to a *particular* solution \vec{V}_0 , while the *general* solution to the problem always remains of the form

$$\vec{V} = \vec{V}_0 + \mathbf{Z}_{\emptyset}^T \cdot \vec{Y}, \quad (64)$$

where \mathbf{Z}_{\emptyset}^T is the transpose of the matrix of the null space of \mathbf{Z} , which is such that $\mathbf{Z}\mathbf{Z}_{\emptyset}^T = 0$, and \vec{Y} is an arbitrary vector of length d_{\emptyset} .

The fact that all applications to date have used the CE in the basis $\mu=0$, requiring an infinite number of CE (or DFT) coefficients to describe the random component of the energy of formation, should have essentially forced the use of a large set of clusters in order to obtain reasonable fitting errors. The need to include large sets of clusters in the CE normally results in an overdetermined fitting problem. Among the most commonly used algorithms to select a particular solution \vec{V}_0 is to assume a certain decay of the expansion coefficients with cluster size or to minimize a cross-validation score for a set of test structures that are not explicitly included in the least-square algorithm. Assuming a specific decay of the expansion coefficients is particularly dangerous since it imposes unknown properties on the function being expanded. At first sight, minimization of a cross-validation score would appear to be more promising although one can obviously achieve optimum fitting of the test structures by including them in the least-square fitting procedure from the beginning (the predictive capability of the CE for structures not included in the fitting is, of course, a valuable *a posteriori* validation criterium). The fundamental flaw with these approaches is that all produce a particular solution \vec{V}_0 to the inversion problem while the general solution remains of the form given by Eq. (64).

Thus, the approach proposed here is to use the rapidly convergent MDWT described in Sec. III B which, with a moderate effort devoted to the calculation of the energies \vec{E} , should result in good fitting of the input energies while keeping $d_{\emptyset}=0$ and the least-square minimization unique. We note, however, that the *only* advantage of working with $d_{\emptyset}=0$ is that one can perform the least-square minimization without having to assume additional properties of the expansion coefficients.

The last observation brings to the forefront the cluster-selection criterium used in the expansion since, given a set of linearly independent structures, it is always possible to find a set of cluster that not only makes $d_{\emptyset}=0$ for the corresponding matrix \mathbf{Z} , but also makes the fitting error as small as desired. In fact, for a CE with n structures, it is relatively straightforward to find a minimum set of n clusters that will invert Eq. (60) exactly or, if preferred, with a very small fitting error. This can be accomplished by considering a *sufficiently* large pool of clusters that will result in zero (or very

small) fitting error from the particular solution \vec{V}_0 for the overdetermined linear system given by Eq. (63). As mentioned, the particular solution will be such that the norm of \vec{V}_0 is minimum, which is equivalent to making all components of \vec{V} in the null space of \mathbf{Z} equal to zero. Since the fitting error is invariant to the vectors \vec{Y} in Eq. (64), it is always possible to find a set of vectors \vec{Y} that will vanish a set of expansion coefficients $\{V_{\eta}\}$, of size d_{\emptyset} , while keeping the fitting error unchanged. We have, of course, considerable freedom in the choice of the algorithm used to select which coefficients $\{V_{\eta}\}$ to make equal to zero. However, in the absence of some physical basis for the selection algorithm, the procedure is simply a mathematical device to achieve good fitting. Although it is recognized that exact or very good fitting of the input energies is not necessarily a desirable outcome for a CE (since the input energies themselves have errors), the point of the argument is that a set of clusters chosen by means of some algorithm aimed solely at reducing the fitting error is not necessarily useful or appropriate.

In the absence of any specific insight into the function being expanded, it appears that the most sensible cluster-selection criterium is provided by the fact that for $\mu=0$ the CE corresponds to a DFT. Thus, as one would do with any DFT, the transform should include wave vectors \vec{k} for the smallest frequencies first. In the cluster-expansion terminology this means $|\vec{k}|=0$, which corresponds to the empty cluster, $|\vec{k}|=1$ which corresponds to the point cluster, $|\vec{k}|=2$ which corresponds to pairs, and so on.

The lack of a real-space distance in the selection criterium remains a problem with the CE since, for example, all pairs correspond to $|\vec{k}|=2$. Therefore there is no *a priori* criterium to decide when a three-point cluster becomes more important in the expansion than a long pair. In any event, the available structures for the fitting procedure will set the maximum number of pairs that can be included in the inversion while keeping $d_{\emptyset}=0$ and, normally, all such pairs should be included before considering higher-order clusters. In general, a limited number of higher-order clusters can be found that also keep the dimension of the null space of \mathbf{Z} equal to zero and, in principle, one can assess their relative contribution to the expansion. In the next section, we will consider different applications of the CE and briefly illustrate the effect of cluster selection under different conditions.

VII. EXAMPLES

In this section we considered a few numerical examples of the inversion of the CE. To implement the inversion, we will consider 68 structures of a body-centered-cubic (bcc) lattice. These 68 structures and the *ab initio* calculation of the energy of formation for the Fe-Co system are described in detail in Ref. 13.

A. Inversion of x^2

The first example we consider is the DFT, or CE for $\mu=0$, of the square of the concentration, i.e., the function $f^{(2)}(\vec{\sigma}) = x^2(\vec{\sigma})$ defined in Sec. III A. The main purpose of this

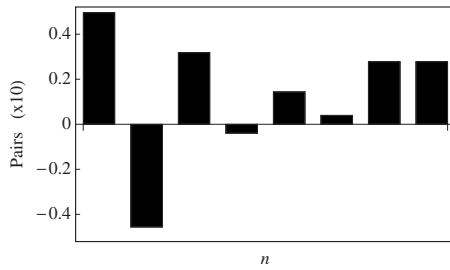


FIG. 1. Coefficients $J_\alpha(0)$ for eight pairs in the CE expansion of the function $f^{(2)}(\vec{\sigma})=x^2(\vec{\sigma})$. The overall fitting error is $\Delta(x^2)_{rms}=0.03$.

example is to illustrate the effect of truncating the CE in a case for which we know that in the limit $N \rightarrow \infty$ an accurate description of the function involves a sum over an infinite number of expansion coefficients that are infinitesimally small. The example is relevant due to the simple fact that we expect the energy of formation of the random alloy to include, in all cases, such a concentration dependence.

Including the first eight pairs in the bcc lattice (plus the empty and point clusters), the matrix \mathbf{Z} in the basis $\mu=0$ for the 68 ordered structures being considered is such that its null space is of dimension $d_\emptyset=0$. Therefore, the solution to the least-square fitting of Eq. (60) is unique. The resulting expansion coefficients $J_\alpha(0)(\times 10)$ for the first eight pairs are shown in Fig. 1. The root-mean-square error of the fit is a very acceptable $\Delta(x^2)_{rms}=0.03$ but we have obviously brought into the description of the function $f^{(2)}(\vec{\sigma})$ a set of spurious pair interactions. More importantly, in a real alloy study, the poor implementation of the CE will imply a fictitious structure of the ordering energy (in this case, for example, a negative effective pair interaction for second neighbors that is not real).

The fitting error of $f^{(2)}(\vec{\sigma})$ can be made as small as desired by increasing the number of clusters used in the expansion. This, of course, requires that we settle on the algorithm used to invert the linear problem of Eq. (60). In the case of $f^{(2)}(\vec{\sigma})$ we have precise knowledge of the function and, in particular, know that the sum of the squares of the expansion coefficients is minimum. Imposing this conditions on the expansion coefficients results in a well-posed least-square fitting problem. Using, arbitrarily, the first 200 pairs in the expansion leads to a fitting of $f^{(2)}(\vec{\sigma})$ with a root-mean-square error

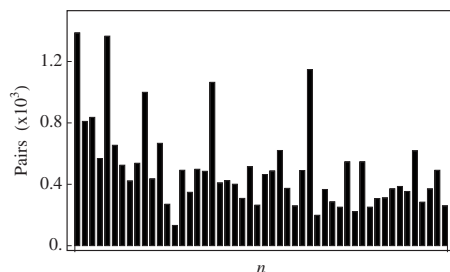


FIG. 2. Coefficients $J_\alpha(0)$ for the first 50 pairs in the cluster expansion of the function $f^{(2)}(\vec{\sigma})=x^2(\vec{\sigma})$. The fitting has been carried out using 200 pairs, in addition to the empty and point clusters, with the condition that the sum of the squares of $J_\alpha(0)$ be minimum. $\Delta(x^2)_{rms}=0$.

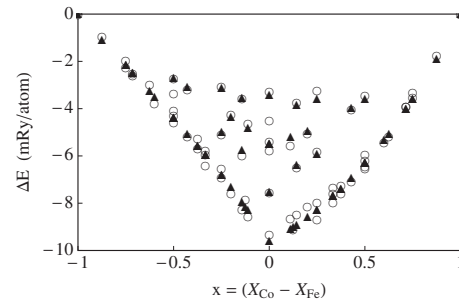


FIG. 3. Energies of formation for 68 Fe-Co compounds obtained by first-principles total-energy calculations in Ref. 13 (open circles), compared to the energies obtained using the VBCE with nn pairs (filled triangles). X_{Fe} and X_{Co} are the atomic concentrations of Fe and Co, respectively. The root-mean-square error is $\Delta E_{rms}=0.27$ mRy/atom.

equal to 0. The expansion coefficients for the first 50 pairs ($\times 10^3$) are shown in Fig. 2. While the inversion is approaching the true DFT of $f^{(2)}(\vec{\sigma})$, in practice, the expansion still generates spurious expansion coefficients that, at best, could be interpreted as “noise” in the CE.

This simple example reinforces our previous observation that attempting to fit the energy of a *random* alloy with a finite number of effective pair and multisite interactions will most likely introduce a nonexistent structure into the *ordering* energy which, itself, may lead to erroneous or dubious predictions regarding the physical behavior of the system. The example also serves to clearly distinguish between the CE, which is a rigorous mathematical representation of functions in configurational space, from the algorithm, assumptions and/or approximations that one might use in the fitting procedure.

B. Fe-Co alloys

As an application to a real alloy system, we carry out the CE and inversion of Eq. (60) using the energies of formation for 68 Fe-Co ordered compounds calculated using a mixed-basis plane-wave pseudopotential method as reported in Ref. 13. We first consider the description of the energies of formation using the VBCE with only two terms in the expansion, namely, the coefficient $J_0(x)$ for the energy of the random alloy and a concentration-dependent nearest-neighbor (nn) pair interaction $J_{\{2,1\}}(x)$. We note that in the VBCE the

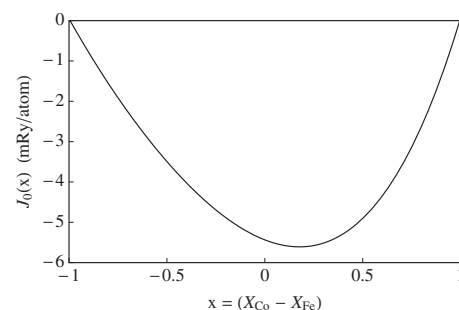


FIG. 4. Coefficient $J_0(x)$ for the VBCE of Fe-Co alloys using nearest-neighbor pairs. $J_0(x)$ is the energy of the random alloy.

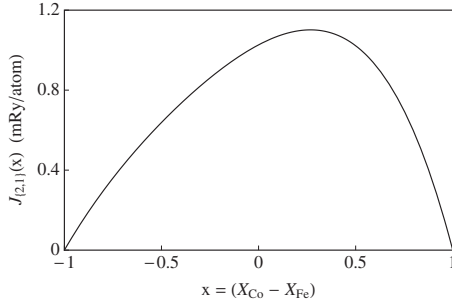


FIG. 5. Nearest-neighbor pair interaction $J_{[2,1]}(x)$ as a function of x for the VBCE of Fe- Co alloys.

basis function $z_1^\mu(\vec{\sigma})$ [see Eq. (49)] vanishes identically since $\tanh(\mu)=x$. Thus, the expansion coefficient $J_1(x)$, which plays the role of a chemical potential, does not enter into the expansion. The next parametrization that needs to be made is to fix the integers $m_\eta^{(0)}$ introduced in Eq. (54). We choose $m_0^{(0)}=4$ and $m_{[2,1]}^{(0)}=2$ for $J_0(x)$ and $J_{[2,1]}(x)$, respectively. With this choice, the total number of fitting parameters $K_\eta^{(m)}$ is eight. Therefore, the energy of formation of the random alloy is described by a polynomial in x of order 4, and the nn effective pair interaction by a polynomial of order 2, times the basis normalization factor $(1-x^2)$ [see Eq. (54)]. In the terminology introduced in Sec. III B, we are including contributions to both $J_0(x)$ and $J_{[2,1]}(x)$ from *all* clusters in the lattice up to four points (included). We also note that for this minimum set of clusters the matrix \mathbf{Z} in Eq. (60) is such that the dimension of its null space is zero. Thus, the inversion of Eq. (60) is unique and is given by the pseudoinverse of \mathbf{Z} , namely, \mathbf{Z}^+ .

In Fig. 3 we compare the energies calculated from first principles to those fitted by means of the VBCE. The overall fitting, as measured by the root-mean-square error, is 0.27 mRy/atom which, although somewhat larger than the relative errors expected from the first-principles calculations, it is quite satisfactory given the fact that we are using the VBCE at the lowest possible level of approximation (i.e., with nn pairs). The coefficients of the expansion as function of concentration are shown in Figs. 4 and 5.

As a final example we consider the VBCE for Fe-Co using a set of clusters consisting of 60 pairs, 20 triangles and

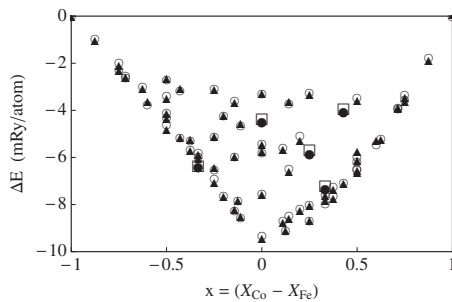


FIG. 6. Energies of formation for 68 Fe-Co compounds obtained by first-principles total-energy calculations (circles) compared to the energies obtained using the VBCE (filled triangles) with 60 pairs, 20 triangles and 10 tetrahedra. The filled circles are five test compound energies not included in the fitting and the empty squares are the energies predicted by the VBCE.

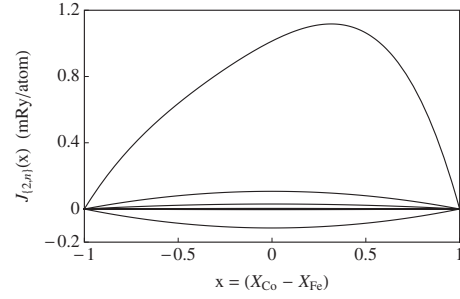


FIG. 7. Pair interactions $J_{[2,n]}(x)$ as a function of x for 60 pairs.

10 tetrahedra. The clusters are selected using a simple compactness criterium based on the sum of the bond lengths. As in our previous example, we use $m_0^{(0)}=4$ for the empty cluster and $m_{[2,1]}^{(0)}=2$ for the nn pairs. For all other clusters, we choose $m_\eta^{(0)}=0$ so that the corresponding coefficients $J_\eta(x)$ depend on concentration only through the basis renormalization factor $(1-x^2)^{n_\eta/2}$. For the set of fitting parameters $\{K_\eta^{(m)}\}$ being considered (91 in total), the dimension of the null space of the matrix \mathbf{Z} is larger than zero and, therefore, the inversion of Eq. (60) has an infinite number of solutions. We approach this problem by building an *approximate* solution to the inversion by successively adding clusters to the expansion to fit the residual error of the previous iteration so that, at each step, the condition $d_\emptyset=0$ holds. The starting point of the fitting procedure is with a set of five pairs, for which $d_\emptyset=0$. Furthermore, we exclude from the fitting five structures chosen at random and use them to test the predictive capability of the method.

A typical result of the fitting is shown in Fig. 6. The root-mean-square error for the 63 structures used in the fitting is $\Delta E_{rms}=0.08$ mRy/atom and that for the five test structures is 0.1 mRy/atom. The very rapid decay of higher-order clusters can be seen in Figs. 7–9, which depict the expansion coefficients for pairs, triangles and tetrahedra. As mentioned, the concentration dependence seen for the expansion coefficients with $m_\eta^{(0)}=0$ comes from the factor $(1-x^2)^{n_\eta/2}$ in Eq. (54). This concentration dependence is simply the normalization factor (or wavelet stretching factor) of the basis functions $z_\eta^\mu(\vec{\sigma})$ when $\tanh(\mu)=x$. Thus, a transformation of the basis sets $\{z_\eta^\mu(\vec{\sigma})\}$ to $\{z_\eta^0(\vec{\sigma})\}$ will cast the piece of the energy of formation described by these coefficients in the form of an Ising-type model with concentration-independent interactions.

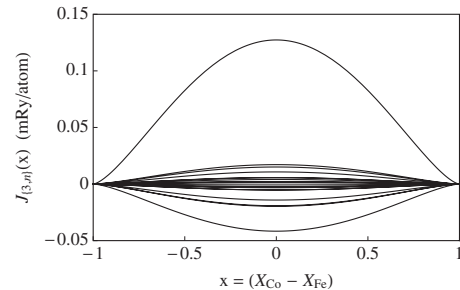


FIG. 8. Three-body interactions $J_{[3,n]}(x)$ as a function of x for 20 triangles.

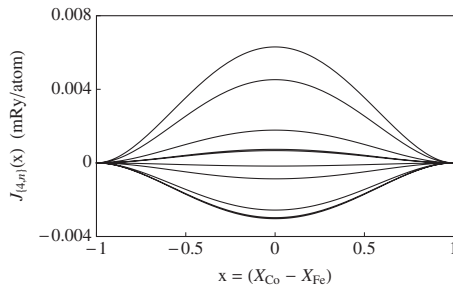


FIG. 9. Four-body interactions $J_{\{4,m\}}(x)$ as a function of x for 10 tetrahedra.

VIII. CONCLUSIONS

In this paper we have developed a rigorous mathematical foundation for the cluster-expansion method. It is shown that the basis introduced by Sanchez *et al.*⁷ corresponds, exactly, to a multidimensional discrete Fourier transform which provides a representation of functions of configuration in terms of their cluster spectrum. The general formalism introduced by Sanchez¹⁰ in 1993 is shown to be a multidimensional discrete wavelet transform that allows to resolve functions of configurations in terms of clusters and concentrations. This wavelet transform, called here a VBCE, results in a rapidly convergent expansion. From the formal point of view, the rigorous mathematical formulation of the CE should settle in a definitive way the criticisms¹¹ raised in the scientific literature regarding the method.

The question of whether or not the energy of formation of alloys can be described by an Ising-type model has a clear and simple answer within the framework of the CE. The VBCE expansion naturally separates the energy of formation into two terms: the energy of formation of the random alloy and the so-called ordering energy. The former depends nonlinearly on concentration while the latter determines the alloy's state of short- and long-range order that might be observed or calculated at given temperature. We have shown that the description of the energy of formation of the random alloy using the CE with $\mu=0$ is a multidimensional DFT that converges slowly and, in principle, requires an infinite number of cluster coefficients for an accurate representation of the energy. Thus, although the CE for $\mu=0$ (or DFT) is certainly a valid representation for any square-integrable function of configuration, for the energy of the random alloy, it does not reduce to an Ising-type model. On the other hand, the ordering energy can be written as rapidly convergent CE with a finite number of terms.

The main advantage of the VBCE is that it provides both the energy of formation of the random state as a function of concentration as well as the ordering energy in the form of a finite and rapidly convergent cluster expansion with concentration-dependent interactions. An alternative approach to address the slow convergence of the DFT for the random state would be to tackle the calculation of the ordering energy directly from first principles. This could be accomplished, for example, by combining electronic-structure methods for the energy of the random state, such as the coherent-potential approximation, with *ab initio* total-energy

calculations for ordered compounds. The first-principles ordering energy could then be cast in the form of an Ising-type model by means of the CE for $\mu=0$ or multidimensional DFT. However, this approach is expected to offer significant challenges since it would be important to treat very different conditions, namely, random state vs perfectly ordered compounds under a set of compatible approximations for the electronic-structure calculations.

In the case of the CE, a practical obstacle is the somewhat limited number of energies normally available to accurately carry out the inversion of Eq. (60). Although this obstacle can be removed with sufficient computational effort, in practice, most applications of the CE to date have been carried out under less than optimum conditions. Consistently, the common approach has been to use the CE expansion in the basis $\mu=0$ which, at least for the energy of formation of alloys, does not converge to a finite set of expansion coefficients in the limit $N \rightarrow \infty$. Combined with a relatively small number of input structures, the practice has forced the introduction into the inversion problem a number of elements that are not necessarily related to the CE formalism and/or to known properties of the function being expanded. These *ad hoc* elements include cluster-selection criteria, assumed specific decays for the cluster interactions in real space and minimization of cross-validation scores.

Of particular note in the context of the theory presented here is the implementation of the CE introduced by Laks *et al.*¹⁴ and subsequently used in CE studies of the energy of formation of $\text{Ga}_{1-c}\text{In}_c\text{P}$ and MoTa alloys.^{15,16} These authors addressed the slow or, as argued here, lack of convergence of the CE for $\mu=0$ by introducing into the energy of formation the sum, calculated in reciprocal space, of all pairs interactions¹⁵ or, alternatively, a “constituent strain energy” which also takes the form of an infinite sum or integral in reciprocal space.¹⁶ We note that the main effect of these infinite reciprocal space sums is to introduce a quadratic dependence in the concentration x into the energy of formation (see, e.g., Fig. 5 in Ref. 16). Therefore, these approximations, which apparently lead to well-converged expansions, are effective devices to partially resolve the nonlinear concentration dependence of the energy of formation brought into evidence by the VBCE.

We have also emphasized the differences between the CE and the assumptions and approximations used for cluster selection and inversion of the linear problem in Eq. (60). The issue of cluster selection is partially resolved by the fact that the CE for $\mu=0$ is a multidimensional DFT. The inclusion of low frequencies, or small values of $|\vec{k}|$, would normally capture the most prominent features of the function being transformed. In the case of the CE, $|\vec{k}|=n$ corresponds to clusters of n points and, therefore, the natural hierarchy of cluster to be included are empty, point, pairs, etc. Unfortunately, the lack of a real-space distance in cluster space does not allow us to distinguish, for example, between different type of pairs since both will correspond to DFT coefficients with $|\vec{k}|=2$. For physical quantities, such as the energy of alloys, our intuition would naturally point us to include compact clusters first. However, the very concept of cluster selection

needs to be approached carefully since it may be influenced by the somewhat arbitrary set of input structures used in the inversion of Eq. (60). If the cluster-selection criterium is guided only by the quality of the fit to the input structures, then it should be considered suspect since it is likely to be a statement on the choice of input structures rather than on the properties of the unknown function we are trying to fit.

We conclude by noting that the result of truncating the CE for $\mu=0$, which nevertheless typically provides a good fit to

the input energies, inevitably introduces a spurious and fictitious structure into the ordering energy, which is key to any statistical thermodynamics theory of alloys. In particular, the ordering energy dictates the state of short- and long-range order in the system and is the basic physical quantity used to understand and, presumably, to predict the physical properties of alloys. It is expected that with the VBCE developed here some of the shortcomings of standard applications of CE method will be satisfactorily addressed.

¹D. de Fontaine, *Solid State Physics* (Academic Press, New York, 1994), Vol. 47, p. 33.

²J. M. Sanchez and D. de Fontaine, *Phys. Rev. B* **17**, 2926 (1978).

³J. M. Sanchez and D. de Fontaine, *Phys. Rev. B* **21**, 216 (1980).

⁴J. M. Sanchez and D. de Fontaine, *Phys. Rev. B* **25**, 1759 (1982).

⁵R. Kikuchi, *Phys. Rev.* **81**, 988 (1951).

⁶J. W. D. Connolly and A. R. Williams, *Phys. Rev. B* **27**, 5169 (1983).

⁷J. M. Sanchez, F. Ducastelle, and D. Gratias, *Physica A* **128**, 334 (1984).

⁸M. Asta, C. Wolverton, D. de Fontaine, and H. Dreyssé, *Phys. Rev. B* **44**, 4907 (1991).

⁹C. Wolverton, M. Asta, H. Dreyssé, and D. de Fontaine, *Phys.*

Rev. B **44**, 4914 (1991).

¹⁰J. M. Sanchez, *Phys. Rev. B* **48**, 14013 (1993).

¹¹A. Gonis, P. P. Singh, P. E. A. Turchi, and X. G. Zhang, *Phys. Rev. B* **51**, 2122 (1995).

¹²R. K. R. Yarlagadda and J. E. Hershey, *Hadamard Matrix Analysis and Synthesis With Applications to Communications and Signal/Image Processing* (Kluwer Academic, Boston, MA, 1997).

¹³A. Díaz-Ortiz, R. Drautz, M. Fähnle, H. Dosch, and J. M. Sanchez, *Phys. Rev. B* **73**, 224208 (2006).

¹⁴D. B. Laks, L. G. Ferreira, S. Froyen, and A. Zunger, *Phys. Rev. B* **46**, 12587 (1992).

¹⁵C. Wolverton and A. Zunger, *Phys. Rev. Lett.* **75**, 3162 (1995).

¹⁶V. Blum and A. Zunger, *Phys. Rev. B* **70**, 155108 (2004).