

Scalable parallel Monte Carlo algorithm for atomistic simulations of precipitation in alloysBabak Sadigh,^{1,*} Paul Erhart,^{1,2,†} Alexander Stukowski,¹ Alfredo Caro,^{1,3} Enrique Martinez,^{1,3} and Luis Zepeda-Ruiz¹¹*Lawrence Livermore National Laboratory, Condensed Matter and Materials Division, Livermore, California, USA*²*Chalmers University of Technology, Department of Applied Physics, Gothenburg, Sweden*³*Los Alamos National Laboratory, Los Alamos, New Mexico, USA*

(Received 21 December 2010; revised manuscript received 17 March 2012; published 11 May 2012)

We present an extension of the semi-grand-canonical (SGC) ensemble that we refer to as the variance-constrained semi-grand-canonical (VC-SGC) ensemble. It allows for transmutation Monte Carlo simulations of multicomponent systems in multiphase regions of the phase diagram and lends itself to scalable simulations on massively parallel platforms. By combining transmutation moves with molecular dynamics steps, structural relaxations and thermal vibrations in realistic alloys can be taken into account. In this way, we construct a robust and efficient simulation technique that is ideally suited for large-scale simulations of precipitation in multicomponent systems in the presence of structural disorder. To illustrate the algorithm introduced in this work, we study the precipitation of Cu in nanocrystalline Fe.

DOI: [10.1103/PhysRevB.85.184203](https://doi.org/10.1103/PhysRevB.85.184203)

PACS number(s): 05.10.Ln, 02.70.Tt, 81.07.Bc, 81.30.Mh

I. INTRODUCTION

The interplay between chemistry and structure is of paramount importance in materials science. This applies in particular to alloys where chemical ordering and precipitation in the presence of surfaces, grain boundaries, dislocations and other structural features lead to complex behavior. Some examples of practical importance include Al-Cu alloys, Ni-Co superalloys as well as steels, the properties of which vary over a wide range depending on composition and microstructure. Understanding and eventually controlling these effects is a prerequisite for designing and improving materials. In principle, modeling and simulation are ideally suited to complement and guide experimental efforts, especially as dimensions shrink and chemical complexity increases.

The objective of the present work is to develop an algorithm that enables us to model the equilibrium properties of phase segregated multicomponent systems containing millions of particles while taking into account chemical degrees of freedom, structural relaxations as well as thermal vibrations. For such an algorithm to be useful on current computing platforms, it must lend itself to efficient parallelization. This is difficult to achieve for Monte Carlo (MC) algorithms that are based on the canonical ensemble.¹ Simulations within the semi-grand-canonical (SGC) ensemble, on the other hand, are easily parallelized but cannot be used to study precipitation and interface formation. The objective of the present work is to develop a MC technique that can both handle multiphase systems and be parallelized easily and efficiently. Note that the parallel algorithm discussed in this paper is suitable for short-range interatomic potentials as described, e.g., by embedded-atom method,² bond-order,³ or Stillinger-Weber⁴ type potentials.

The paper is organized as follows. In Sec. II, we discuss how to model chemical mixing and phase segregation on the atomic scale. The most common approach is to sample the chemical configuration space using transmutational MC methods, which require as key ingredient an appropriate statistical ensemble. Following a discussion of the advantages and shortcomings of existing ensembles with respect to the present application,

we introduce the variance-constrained semi-grand-canonical (VC-SGC) ensemble, which can be viewed as a generalization of the extended Gaussian ensemble technique to multicomponent systems,^{5,6} and formulate a simple serial VC-SGC-MC algorithm. In Sec. III, we address the question how the MC methods introduced in Sec. II can be adapted for simulations of systems containing millions of particles. To this end, we derive transition matrices and their efficient decomposition. In Sec. IV, we finally discuss the simultaneous and efficient sampling of chemical, structural and vibrational degrees of freedom, and consider the precipitation of Cu in nanocrystalline Fe as an illustrative example.

The algorithms developed in this work have been implemented in the massively parallel molecular dynamics code LAMMPS⁷ and the source code is available from the authors.

II. MODELING CHEMICAL MIXING AND PRECIPITATION

On the atomic scale, chemical mixing in alloys is most commonly studied using MC simulations within either the SGC or the canonical ensemble. Therefore we first discuss in some detail these two ensembles before deriving the VC-SGC ensemble, which merges the advantages of the canonical and SGC ensembles. In the following, we use the subscripts C, S, and V to indicate quantities that are connected to the canonical, SGC and VC-SGC ensembles, respectively. For the sake of simplicity, we limit our discussion to binary alloys. The generalization to systems containing an arbitrary number of species is straightforward.

Consider a system of N particles confined in a box of volume V , where each particle carries a spin of value 0 or 1. A configuration of this system can be denoted $(\mathbf{x}^{3N}, \sigma^N)$, where \mathbf{x}^{3N} is a $3N$ -dimensional vector describing the positions of every particle, and σ^N is an N -dimensional spin vector. The number of spin 1 particles is $n = \sum_{i=1}^N \sigma_i$, and their concentration $c = n/N$. We denote the energy of a configuration by $U(\mathbf{x}^{3N}, \sigma^N)$.

A. The canonical ensemble

The canonical ensemble describes the thermodynamics of systems that are chemically isolated, i.e., the number of members of each species is kept constant. The partition function for the canonical ensemble at temperature T for the binary system defined above is

$$\mathcal{Z}_C(c, \mathcal{N}) = \Lambda_1^{-3(N-n)} \Lambda_2^{-3n} \frac{1}{n!(N-n)!} \times \int \exp[-\beta U(\mathbf{x}^{3N}, \sigma^N)] d^{3N} \mathbf{x}, \quad (1)$$

where $\beta = 1/k_B T$, $\Lambda_i = \sqrt{h^2/2\pi m_i k_B T}$ is the thermal de Broglie wavelength for component i , and $\mathcal{N} = \{N, V, T\}$ is the set of independent thermodynamic variables.¹ Monte Carlo simulations in this ensemble sample the probability distribution

$$\pi_C(\mathbf{x}^{3N}, \sigma^N; c, \mathcal{N}) \propto \exp[-\beta U(\mathbf{x}^{3N}, \sigma^N)]. \quad (2)$$

Efficient sampling of the above distribution involves two kinds of trial moves: (i) particle displacements $\mathbf{x}^{3N} \rightarrow \mathbf{x}_i^{3N}$ and (ii) compositional changes $\sigma^N \rightarrow \sigma_i^N$ that keep the concentration fixed. In practice, in trial move (i) a particle is selected at random and assigned a random displacement, while for trial move (ii) two particles with unlike spins are selected at random and their spins are exchanged. These trial moves are accepted with probability

$$\mathcal{A}_C = \min\{1, \exp[-\beta \Delta U]\}, \quad (3)$$

$$\Delta U = U(\mathbf{x}_i^{3N}, \sigma_i^N) - U(\mathbf{x}^{3N}, \sigma^N). \quad (4)$$

This acceptance probability is designed to satisfy detailed balance. Approach to equilibrium can be accelerated substantially if trial moves (i) are biased along the force vector $-\nabla U(\mathbf{x}^{3N}, \sigma^N)$. This is achieved most easily via a hybrid technique where particle positions \mathbf{x}^{3N} are sampled via molecular dynamics (MD), while spin degrees of freedom are sampled using the spin exchange (transmutation) MC moves described above.

B. The semigrandcanonical ensemble

The SGC ensemble describes the thermodynamics of a system in contact with an infinite reservoir at constant temperature and chemical potential for each species. This ensemble corresponds to a set of configurations with varying compositions, but with their ensemble average constrained by the reservoir. The equilibrium probability distribution of the SGC ensemble for the binary system defined above thus becomes

$$\begin{aligned} \pi_S(\mathbf{x}^{3N}, \sigma^N; \Delta\mu, \mathcal{N}) \\ \propto \exp\{-\beta[U(\mathbf{x}^{3N}, \sigma^N) + \Delta\mu N \hat{c}(\sigma^N)]\}, \\ \hat{c}(\sigma^N) = \frac{1}{N} \sum_{i=1}^N \sigma_i, \end{aligned} \quad (5)$$

where $\Delta\mu$ is a Lagrange multiplier that constrains the average concentration. The partition function can be expressed in terms

of the canonical one via

$$\mathcal{Z}_S(\Delta\mu, \mathcal{N}) = \int_0^1 \mathcal{Z}_C(c, \mathcal{N}) \exp[-\beta \Delta\mu N c] dc. \quad (6)$$

The SGC ensemble can be sampled using a Monte Carlo algorithm where trial moves $\sigma^N \rightarrow \sigma_i^N$ are made by (i) selecting a particle at random, (ii) flipping its spin, (iii) computing the change in energy ΔU , and concentration Δc . Trial moves are accepted with probability

$$\mathcal{A}_S = \min\{1, \exp[-\beta(\Delta U + \Delta\mu N \Delta c)]\}, \quad (7)$$

which is designed to satisfy detailed balance.

The acceptance probability expression above has important physical significance. It shows that in the SGC ensemble, the force associated with a change in the chemical configuration does not solely originate from the potential energy function ΔU , but also from the term $\Delta\mu N \Delta c$. In particular, for any change in concentration, a constant external chemical driving force $\Delta\mu N$ is added to the usual interatomic forces in order to drive the equilibrium concentration to the desired value. In physical experiments, $\Delta\mu$ corresponds to the chemical potential difference between the two species. In practice, it alters the acceptance probability (7) for trial moves that lead to a concentration change. It is important to note that in this way, only single-phase equilibria can be established. This means that, e.g., for immiscible systems such as the one shown in Fig. 1(a), concentrations inside the miscibility gap cannot be stabilized. This limitation results from the functional

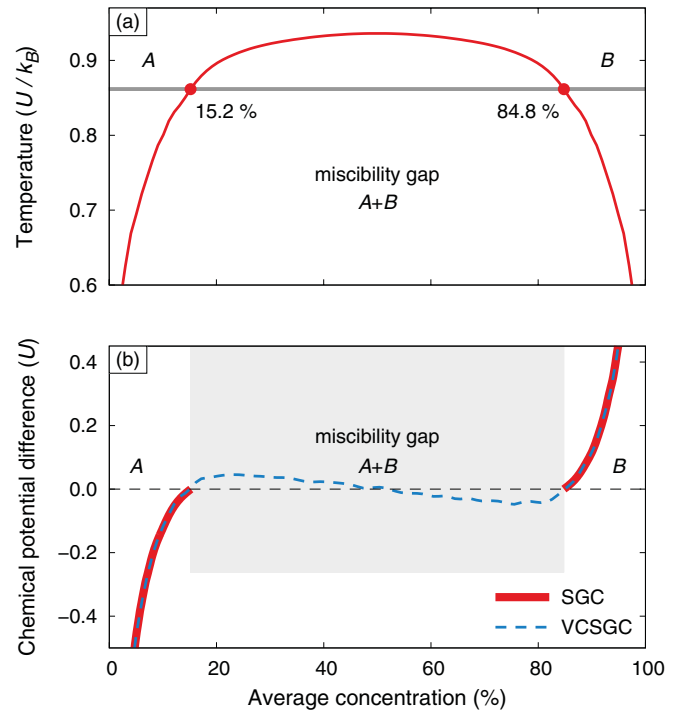


FIG. 1. (Color online) (a) Phase diagram for the Ising-type model system described in the text. The horizontal bar marks the temperature of $0.8617 U/k_B$ at which all the simulations described in this paper have been carried out. The closed circles indicate the solubility limits at this temperature. (b) The chemical driving force $\Delta\mu$ as a function of concentration as obtained from a series of simulations in the SGC (solid line) and VC-SGC (dashed line) ensembles, respectively.

dependence between the chemical potential difference $\Delta\mu$ and average concentration $\langle\hat{c}\rangle_S$ not being one-to-one in the multiphase regions of the phase diagram.

To illustrate this point, let us consider an Ising-type Hamiltonian

$$\mathcal{H} = \frac{1}{2} \sum_{i \in A, j \in A} \epsilon_{AA}(r_{ij}) + \frac{1}{2} \sum_{i \in A, j \in B} \epsilon_{AB}(r_{ij}) + \frac{1}{2} \sum_{i \in B, j \in B} \epsilon_{BB}(r_{ij}), \quad (8)$$

where r_{ij} denotes the neighbor shell of site i in which site j is located. We use a body-centered cubic (BCC) lattice with interactions up to the second-neighbor shell and $\epsilon_{AA}(1) = \epsilon_{BB}(1) = -10U$, $\epsilon_{AB}(1) = -9.7U$, and $\epsilon_{AA}(2) = \epsilon_{BB}(2) = \epsilon_{AB}(2) = -2U$. The phase diagram for this model system can be calculated analytically and is shown in Fig. 1(a). We carried out a series of simulations using the SGC-MC method for a system containing 2000 sites at a temperature of $0.8617U/k_B$, starting from a solid solution at 50%. The dependence of $\Delta\mu$ on $\langle\hat{c}\rangle_S$ determined in this way is depicted by the solid red line in Fig. 1(b). Note the discontinuity in the $\Delta\mu - \langle\hat{c}\rangle_S$ plot, which occurs in the region of the binary phase diagram where the miscibility gap is located. This demonstrates that the SGC-MC method is not suitable for studying phase segregation.

C. The variance-constrained semigrandcanonical ensemble

To simulate systems in multiphase regions of phase diagram where precipitation occurs, we modify the SGC ensemble by adding a constraint that fixes the ensemble-averaged squared concentration $\langle\hat{c}^2\rangle$. This limits concentration fluctuations and thus, when inside the miscibility gap, prevents the concentration to fluctuate to the phase boundaries. We refer to this approach as the VC-SGC ensemble, which can be categorized as an extended Gaussian ensemble. Such ensembles describe the thermodynamics of systems in contact with finite reservoirs.⁶ We will show below that the VC-SGC ensemble is ideal for studying equilibrium properties of multiphase systems and that it is quite straightforward to devise Monte Carlo algorithms that sample this ensemble.

In contrast to the SGC ensemble that is characterized by an infinite reservoir with constant chemical potential $\Delta\mu$, the reservoir of the VC-SGC ensemble is controlled by two independent parameters ϕ and κ . The statistical mechanical origin of these parameters is laid out in detail in Appendix. There it is shown that ϕ and κ are Lagrange multipliers associated with constraints on the first and the second moments of the concentration, respectively. The most probable distribution subject to these constraints is then derived to be [see Eq. (A4)]

$$\pi_V(\mathbf{x}^{3N}, \sigma^N; \phi, \kappa, \mathcal{N}) \propto \exp[-\beta U(\mathbf{x}^{3N}, \sigma^N)] \times \exp\{-\beta N \hat{c}(\sigma^N) [\phi + \kappa N \hat{c}(\sigma^N)]\}. \quad (9)$$

We can thus express the partition function of the VC-SGC ensemble in terms of the canonical one as

$$\mathcal{Z}_V(\phi, \kappa, \mathcal{N}) = \int_0^1 \mathcal{Z}_C(c, \mathcal{N}) \exp[-\beta N c (\phi + \kappa N c)] dc. \quad (10)$$

The VC-SGC ensemble can be considered a generalization of both the SGC and the canonical ensembles. The former is obtained trivially by letting $\kappa \rightarrow 0$. In order to obtain the canonical ensemble, we complete the square in Eq. (9) and rewrite the VC-SGC probability distribution as

$$\pi_V(\mathbf{x}^{3N}, \sigma^N; \phi, \kappa, \mathcal{N}) \propto \exp[-\beta U(\mathbf{x}^{3N}, \sigma^N)] \times \exp\left\{-\beta \kappa \left[N \hat{c}(\sigma^N) + \frac{\phi}{2\kappa} \right]^2\right\}. \quad (11)$$

The canonical ensemble is recovered when $\kappa \rightarrow \infty$ and $\phi = -2\kappa N c$. This can be seen by rewriting the canonical partition function as

$$\mathcal{Z}_C(c, \mathcal{N}) = \int_0^1 \mathcal{Z}_C(c', \mathcal{N}) \delta(c - c') dc'. \quad (12)$$

Hence the VC-SGC ensemble may be obtained by generalizing the δ function that fixes the concentration in the canonical ensemble to a Gaussian with tunable width determined by the parameter κ . Now in multiphase regions of phase diagrams where the SGC ensemble is not stable, a VC-SGC ensemble can be devised by judiciously choosing the two parameters ϕ and κ that combine both advantages of the SGC and the canonical ensembles. Traditionally, the canonical ensemble has been used to study precipitation inside the miscibility gap. Our objective with this paper is to show that the same physics can be studied much more efficiently in the VC-SGC ensemble, especially when parallel computing is utilized.

Thanks to its similarity with the SGC ensemble, it is straightforward to formulate a MC algorithm for sampling the VC-SGC ensemble, where trial moves $\sigma^N \rightarrow \sigma_t^N$ comprise (i) selecting a particle at random, (ii) flipping its spin, and (iii) computing the change in energy ΔU and concentration Δc as well as

$$\tilde{c} = \frac{\hat{c}(\sigma_t^N)^2 - \hat{c}(\sigma^N)^2}{2\Delta c} = \frac{\hat{c}(\sigma_t^N) + \hat{c}(\sigma^N)}{2}. \quad (13)$$

These trial moves are accepted with probability

$$\mathcal{A}_V = \min(1, \exp\{-\beta[\Delta U + N\Delta c(\phi + 2\kappa N\tilde{c})]\}). \quad (14)$$

Once again, this acceptance probability is designed to satisfy detailed balance. The force associated with a change in spin configuration receives contributions from both the change in the interatomic potential energy function ΔU as well as the external concentration dependent force $N\Delta c(\phi + 2\kappa N\tilde{c})$. Hence, for a change in concentration, the usual interatomic forces are augmented with an additional external chemical driving force that at variance with the SGC ensemble is not a constant but varies linearly with concentration as $N\phi + 2\kappa N^2 c$. When ensemble-averaged, the equilibrium chemical driving force that corresponds to the chemical potential difference in physical experiments and the $\Delta\mu$ parameter in the SGC ensemble now becomes

$$\Delta\mu = \phi + 2\kappa N \langle\hat{c}\rangle_V. \quad (15)$$

This very important relation is derived in the appendix, see Eq. (A11). It connects the VC-SGC and the SGC ensembles and will be used extensively in the following to design and

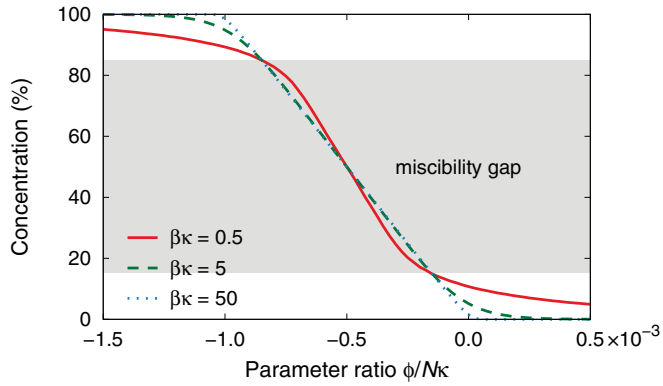


FIG. 2. (Color online) Dependence of global concentration on the parameter ratio $\phi/N\kappa$ obtained from VC-SGC-MC simulations. All simulations were carried out at a temperature of $0.8617 U/k_B$ for the model system described in Sec. II B.

analyze Monte Carlo simulations of systems in which several phases coexist.

We now apply the VC-SGC-MC method to study the model system described in Sec. II B. Figure 2 illustrates the relation between the global concentration and the parameter ratio ϕ/κ . It clearly demonstrates that using the VC-SGC-MC algorithm enables us to stabilize the system at arbitrary global concentrations in and outside the miscibility gap.

The dependence of the standard deviation of the concentration on the variance parameter κ follows a power law [see Fig. 3(a)], $\langle \Delta c^2 \rangle_V \propto 1/\sqrt{\kappa}$. The relation between the

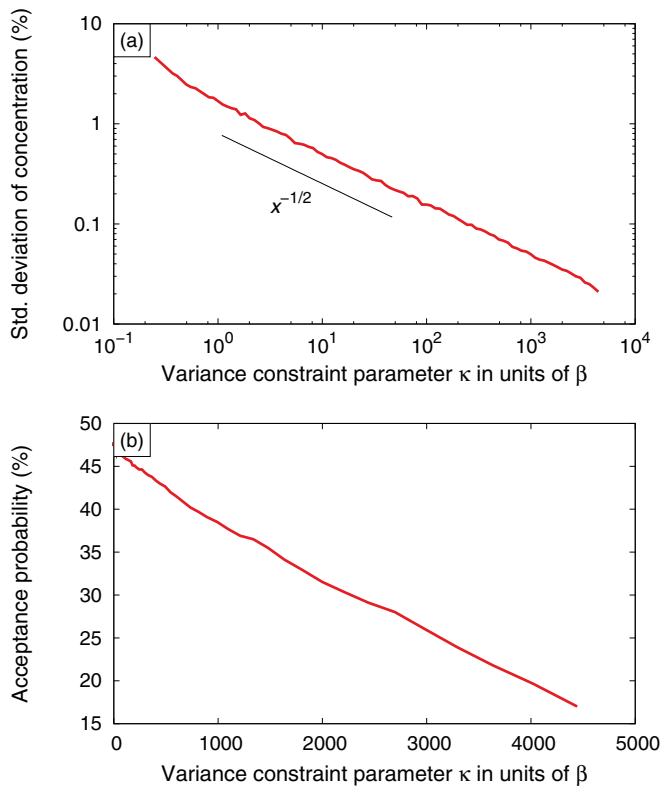


FIG. 3. (Color online) Dependence of (a) standard deviation of concentration and (b) acceptance probability on the variance constraint parameter κ .

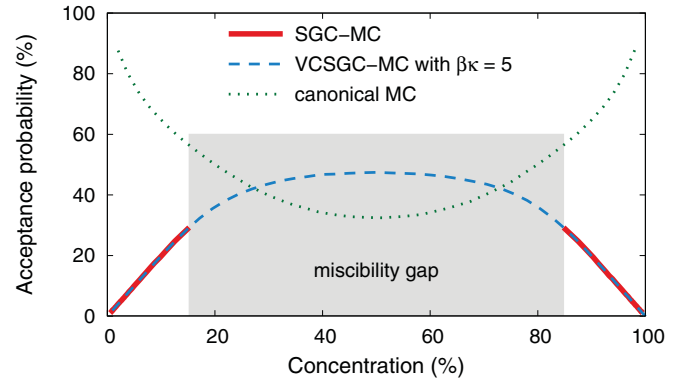


FIG. 4. (Color online) Comparison of the acceptance rate as a function of global concentration for the SGC, VC-SGC ($\kappa = 5$), and canonical MC methods. At low concentrations, the canonical MC method yields the highest acceptance rates while inside the miscibility gap the VC-SGC-MC achieves the best results.

acceptance probability and κ , on the other hand, is linear with a negative slope [see Fig. 3(b)]. Increasing κ thus has two effects: it leads to a smaller standard deviation, while simultaneously reducing the acceptance probability.

We can also compare the acceptance probability as obtained with the VC-SGC-MC method with the results for the SGC and canonical MC methods. As shown in Fig. 4, in the single-phase regions of the phase diagram the SGC and VC-SGC-MC methods coincide and produce comparably low acceptance rates, while the canonical MC method provides large acceptance rates. However, inside the miscibility gap, which is the region of interest when it comes to phase segregation, the VC-SGC method yields the best results.

We now study the functional dependence of the chemical driving force $\Delta\mu$ obtained from Eq. (15) on the average concentration using the VC-SGC-MC method. The result is shown in Fig. 1(b) in comparison with the data obtained using the SGC-MC method. The VC-SGC-MC method produces a continuous relation between $\Delta\mu$ and c throughout the entire concentration range. In the single-phase regions of the phase diagram, the SGC and VC-SGC-MC results coincide. Inside the miscibility gap where the SGC-MC fails, the VC-SGC-MC method reproduces the van-der-Waals loop associated with the formation of phase boundaries.⁸ This is a very important result that can be used to extract interface free energies.⁹

To summarize, the VC-SGC-MC method imposes a constraint on the variance of the concentration, and allows for equilibration at arbitrary global concentrations. Thereby, it merges the advantages of the SGC and the canonical MC algorithms. In the next section, we show that the VC-SGC-MC algorithm is also very well suited for parallelization enabling simulations of systems with many million particles.

III. PARALLELIZATION STRATEGIES FOR LARGE SYSTEMS

There are a multitude of problems involving precipitation, especially in the presence of structural defects such as dislocations, grain boundaries, and surfaces, which require simulations of systems with hundreds of thousands or millions of particles. Efficient parallelization schemes with good

scalability are a necessity in order to address these problems. Here, we focus on short-range interaction potentials as described, e.g., by embedded-atom method,² bond-order,³ or Stillinger-Weber⁴ type potentials.

Monte Carlo simulations in the canonical ensemble do not lend themselves to efficient parallelization since trial moves in this scheme involve exchange of two particles that can be located on any two processors. Although it is possible to conceive elaborate distributed algorithms, it is difficult to implement a scheme that ensures unbiased sampling and still avoids spending a considerable fraction of simulation time on interprocessor communication. The SGC ensemble, on the other hand, can be parallelized easily but, as discussed in Sec. II B, cannot be used to study precipitation. The purpose of this work is to develop a Monte Carlo technique that can both handle multiphase systems and can be parallelized easily and efficiently. In the following, we discuss parallelization strategies for the SGC as well as the VC-SGC ensembles and demonstrate their excellent scalability and efficiency.

A. Domain decomposition for sampling trial moves

Consider for simplicity a simulation box in the shape of a cube with linear dimension L . In systems with short-range interactions, the most common parallelization strategy is to subdivide the simulation box into a regular lattice of N_{CPU} equivalent cells $\{C_i\}$ with linear dimension $L_c = L/N_c$, where $N_{\text{CPU}} = N_c \times N_c \times N_c$. (The generalization to noncubic cells is straightforward).

At every Monte Carlo step, a cubic domain \mathcal{D}_i is chosen inside each cell C_i in such a way as to ensure that equivalent domains on different processors are *noninteracting*. This means that the total energy change ΔU associated with arbitrary spin flips inside the domains $\{\mathcal{D}_i\}$ can be written as the sum of the *independent* local energy changes ΔU_i on each processor, i.e., $\Delta U = \sum_{i=1}^{N_{\text{CPU}}} \Delta U_i$. Note that all domains \mathcal{D}_i are equivalent with linear dimension $L_D = L_c - R_c$, where R_c is the effective interaction radius in the system. For pair interactions, this radius equals the cutoff radius of the potential, while for three-body potentials, it is usually twice the cutoff radius.

It is easy to see that for the above parallelization strategy to be possible the linear dimension L_c must be larger than R_c . Let us first discuss the case when L_c is exactly twice R_c . In this case, the independent domains will have the linear dimension $L_D = R_c$. They constitute the eight nonoverlapping octants of each cell C_i as depicted in Fig. 5. In this figure, all domains “A” are noninteracting and so are all domains “B,” etc. At each Monte Carlo trial move, one of the eight octants is chosen at random. It is important that all cells C_i work on the same octant simultaneously since only in this way, the trial moves on different processors are with certainty noninteracting.

The above method of subdividing each cell C_i into eight non-overlapping octants also works when $L_c > 2R_c$. However, bear in mind that confining the local trial moves to noninteracting domains produces weak spatial correlations that can slow down the approach to equilibrium, especially when phase segregation and growth of precipitates is expected. These spatial correlations are minimized if the total volume of the domains $\{\mathcal{D}_i\}$ is maximized. This can be achieved

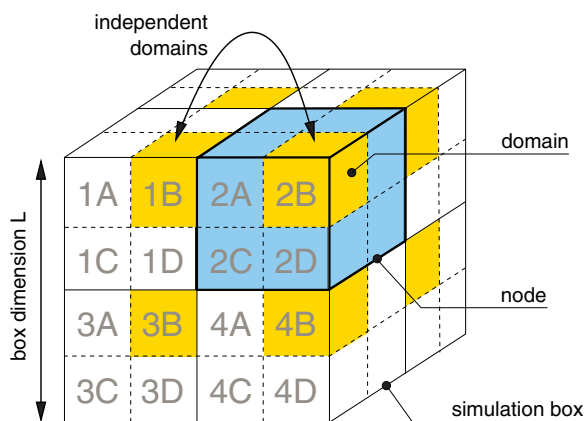


FIG. 5. (Color online) Spatial decomposition (solid lines) and subsequent division into octants (dashed lines) of a system with short-ranged interactions. Sets of octants with the same letter are independent of each other. One such set is marked in yellow.

by growing each octant to a cube with linear dimension $L_D = L_c - R_c$. The eight distinct domains thus generated inside each cell C_i do overlap. This leads to the central region of C_i be covered by all eight \mathcal{D}_i . To ensure uniform sampling, the particles in the outer regions of the C_i cells must be selected with higher probability than those in the center. This can be achieved by assigning differential weights to the particles in the system depending on their position inside C_i (see the right panel of Fig. 6) prior to making trial moves.

It is now straightforward to devise an efficient parallel Monte Carlo algorithm where each trial move is composed of N_{CPU} local moves $\sigma_i^N \rightarrow \sigma_i^N + \Delta\sigma_i^N(i)$ carried out inside the domains $\{\mathcal{D}_i\}$ synchronously on all processors. To ensure uniform sampling, a trial move is constructed in two stages: (i) select one of the eight independent domains $\{\mathcal{D}_i\}$ at random and broadcast to all processors; message passing can be avoided by synchronizing the seed for the random number generator on all processors, and (ii) on each processor i , pick a particle at random inside the chosen domain and flip its spin. Different parts of the domain may be sampled with different weights.

It is important to note that the composite trial move $\sigma^N \rightarrow \sigma^N + \sum_{i=1}^{N_{\text{CPU}}} \Delta\sigma_i^N(i)$ constructed in this way will be rejected at a very high rate. In the following section, we describe how one can improve the above procedure in order to obtain reasonable acceptance probabilities for composite trial moves.

B. Parallel Monte Carlo algorithms

1. Monte Carlo sampling of SGC ensemble

In this section, we describe how one can devise parallel Monte Carlo simulations in the SGC ensemble with composite trial moves constructed from trial moves simultaneously generated on all processors. The algorithm is as follows: (i) on each processor i make a local trial move $\Delta\sigma_i^N(i)$ according to one of the procedures described in Sec. III A and (ii) compute the local changes in energy ΔU_i and concentration Δc_i and accept this move with probability

$$A_S^p(i) = \min\{1, \exp[-\beta(\Delta U_i + \Delta\mu N \Delta c_i)]\}, \quad (16)$$

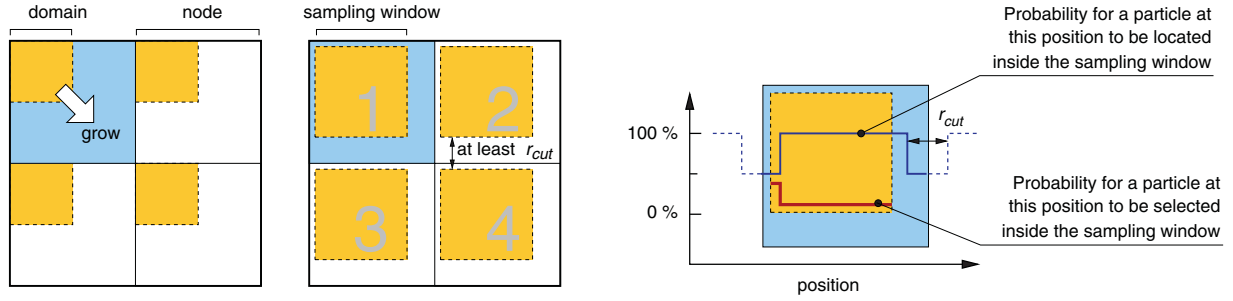


FIG. 6. (Color online) Schematic representation of an optimal spatial decomposition (compare Fig. 5). For a pair potential, the domains have to be separated by at least one cutoff distance, while for potentials with three-body terms, the separation has to be at least two cutoff distances.

otherwise set $\Delta\sigma_i^N(i) = 0$. The global composite trial move is now $\sigma^N \rightarrow \sigma^N + \sum_{i=1}^{N_{\text{CPU}}} \Delta\sigma_i^N(i)$. Thanks to the independence of the domains \mathcal{D}_i , the transition probability for this move is proportional to $\prod_{i=1}^{N_{\text{CPU}}} \mathcal{A}_S^p(i)$ and satisfies detailed balance.

2. Monte Carlo sampling of VC-SGC ensemble

The similarity of the SGC and VC-SGC ensembles discussed in Sec. II C suggests that parallelization strategies might be similar as well. A closer inspection, however, reveals that for a composite trial move $\sigma^N \rightarrow \sigma_i^N$, where $\sigma_i^N = \sigma^N + \sum_{i=1}^{N_{\text{CPU}}} \Delta\sigma_i^N(i)$, we have

$$\hat{c}(\sigma_i^N)^2 - \hat{c}(\sigma^N)^2 \neq \sum_{i=1}^{N_{\text{CPU}}} \hat{c}[\sigma^N + \Delta\sigma_i^N(i)]^2. \quad (17)$$

This implies that there is a coupling between the domains \mathcal{D}_i , and as a result, the simple method outlined in the previous section for the SGC ensemble cannot be directly applied to the parallel sampling of the VC-SGC ensemble. To resolve this issue, we first modify the acceptance probability distribution Eq. (14) for the serial sampling of the VC-SGC ensemble as follows:

$$\mathcal{A}_V = \min(1, \exp[-\beta[\Delta U + N\Delta c(\phi + 2\kappa N c_0)]]) \times \min\{1, \exp[-\beta\kappa N^2 \Delta c(\tilde{c} - c_0)]\}, \quad (18)$$

where \tilde{c} was defined in Eq. (13). It is easy to verify that the acceptance probability distribution in Eq. (18) satisfies detailed balance. The parameter c_0 introduced in Eq. (18) can change the acceptance probability and thus the approach to equilibrium but it does not affect the final result. An optimal choice is

$$c_0 = \langle \hat{c} \rangle_V. \quad (19)$$

In practice, the simulations are performed with c_0 chosen to be the best guess for the average concentration. In Sec. III C, we will explicitly demonstrate the correlation between c_0 , κ , and ϕ and discuss acceptance rates for the simple Ising model introduced earlier.

We can now formulate a parallel Monte Carlo algorithm in the VC-SGC ensemble with composite trial moves comprising N_{CPU} synchronous local moves $\sigma^N \rightarrow \sigma^N + \sum_{i=1}^{N_{\text{CPU}}} \Delta\sigma_i^N(i)$: (i) on each processor i make a local trial move as detailed in the Sec. III A and (ii) compute the local changes in energy ΔU_i

and concentration Δc_i and accept this move with probability

$$\mathcal{A}_V^{p,\text{loc}}(i) = \min(1, \exp\{-\beta[\Delta U_i + N\Delta c_i(\underbrace{\phi + 2\kappa N c_0}_{=\Delta\mu_0})]\}), \quad (20)$$

otherwise set $\Delta\sigma_i^N(i) = 0$. Following Eq. (18), the global trial move $\sigma^N \rightarrow \sigma^N + \sum_{i=1}^{N_{\text{CPU}}} \Delta\sigma_i^N(i)$ may be accepted with probability

$$\begin{aligned} \mathcal{A}_V^{p,\text{glob}} &= \min\left\{1, \exp\left[-2\beta\kappa N^2 \sum_i \Delta c_i(\tilde{c}_i - c_0)\right]\right\} \\ &= \min[1, \exp(-\beta\kappa N^2 \Delta c_{\text{tot}}\{\Delta c_{\text{tot}} - 2[\hat{c}(\sigma^N) - c_0]\})], \end{aligned} \quad (21)$$

where $\Delta c_{\text{tot}} = \sum_{i=1}^{N_{\text{CPU}}} \Delta c_i$ is the total change in concentration due to the composite trial move. This quantity can be efficiently computed using for example the message passing interface¹⁰ altogether command.

C. Efficiency of the parallel VC-SGC-MC method

In arriving at Eq. (20), we have introduced the parameter c_0 and the abbreviation $\Delta\mu_0 = \phi + 2\kappa N c_0$. Together with κ these parameters determine the average and the variance of the concentration. In this section, we will demonstrate the correlation between these parameters using the simple Ising model described in Sec. II B.

The derivation of the transition matrix for the parallel VC-SGC-MC method in the previous section revealed a close resemblance with the parallel SGC-MC method. In particular, the acceptance probabilities $\mathcal{A}_S^p(i)$ and $\mathcal{A}_V^{p,\text{loc}}(i)$ in Eqs. (16) and (20) become identical if $\Delta\mu_0 = \Delta\mu$. This of course requires c_0 to be chosen according to the optimality condition Eq. (19). This insight greatly simplifies the choice of parameters for the parallel VC-SGC-MC method.

In Fig. 7(a), we show the average concentration obtained in simulations with different values of $\Delta\mu_0$ and κ , for a fixed target concentration of $c_0 = 0.25$ located inside the miscibility gap. All simulations were carried out using 64 CPUs, a $4 \times 4 \times 4$ decomposition, and a BCC lattice with $40 \times 40 \times 40$ conventional unit cells. The number of particles per processor is thus equal to the number of particles in the serial VC-SGC-MC simulations discussed in Sec. II C.

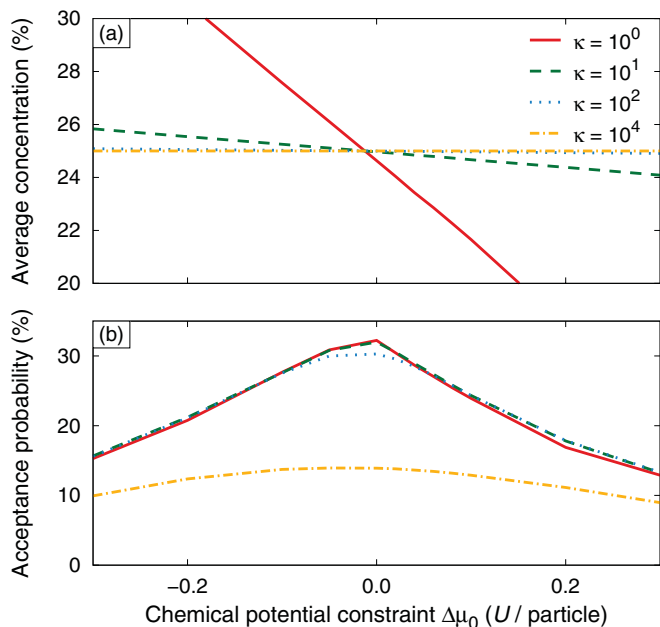


FIG. 7. (Color online) (a) Average concentration and (b) acceptance probability obtained from parallel VC-SGC-MC simulations using different combinations of the parameters $\Delta\mu_0$ and κ for $c_0 = 0.25$ in Eq. (21).

For small values of κ , the average concentration varies strongly with $\Delta\mu_0$. As κ is increased, the total concentration is confined to small variations about c_0 and the average concentration becomes less sensitive to the choice of $\Delta\mu_0$. Comparison with Fig. 1(b), where the chemical driving force is shown as a function of average concentration (\hat{c}), reveals that $\langle \hat{c} \rangle$ equals c_0 exactly when $\Delta\mu = \Delta\mu_0$. This confirms Eq. (19) and validates the underlying connection between the SGC and VC-SGC-MC methods.

While for sufficiently large values of κ , the parameter $\Delta\mu_0$ does not affect the average concentration, it does have a significant impact on the acceptance probability as illustrated in Fig. 7(b). For a given value of κ , the acceptance probability becomes maximal if $\Delta\mu = \Delta\mu_0$, which again confirms the optimality condition Eq. (19). Similar to the case of the serial VC-SGC-MC algorithm [compare Fig. 3(b)], we also find that for a fixed value of $\Delta\mu_0$, the acceptance probability decreases with increasing κ as shown explicitly in Fig. 8. It is, however, remarkable that over a rather wide range, the value of κ does not have a significant negative impact on the acceptance probability.

Now that we have understood the interplay between the parameters $\Delta\mu_0$, κ , and c_0 , we can formulate an optimal strategy for choosing their values: (i) determine the chemical driving force $\Delta\mu_S$ in the vicinity of the two-phase region using the SGC-MC method. This requires simulations involving only small system sizes since we are only interested in single-phase equilibria [compare Fig. 1(b)]. (ii) Choose a value of κ based on the desired standard deviation of the concentration (compare Fig. 3). (iii) Set $\Delta\mu_0 = \Delta\mu_S$ and c_0 to the desired concentration inside the miscibility gap. In this way, the parameter $\phi = \Delta\mu_0 - 2\kappa N c_0$ is determined as well. For all subsequent simulations inside the miscibility gap, $\Delta\mu_0$ can be

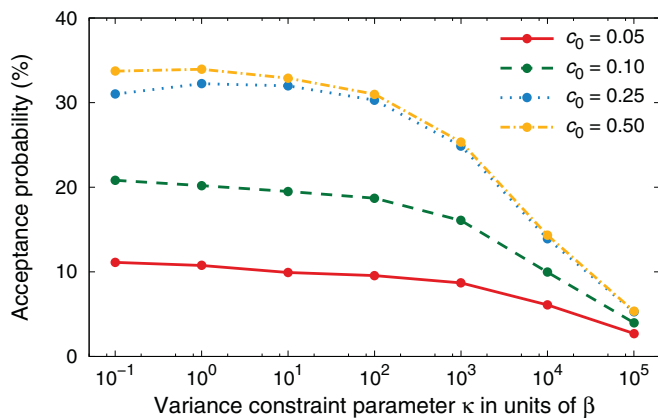


FIG. 8. (Color online) Acceptance probability as a function of variance constraint parameter κ for different target concentrations c_0 and optimal values for $\Delta\mu_0$ as obtained with the parallel VC-SGC-MC algorithm.

held fix and only c_0 is tuned to obtain the desired concentration. From Fig. 8, one observes that at an average concentration of 50% the parallel VC-SGC-MC algorithm achieves a maximal acceptance ratio of about 34%, which compares favorably with a maximum value of about 47% for the serial VC-SGC-MC method (see Fig. 3).

To investigate the performance of the parallel VC-SGC-MC algorithm in the weak-scaling limit, a series of simulations with an increasing number of processors was carried out in which the number of particles per processor was kept constant (2000 particles, $10 \times 10 \times 10$ conventional unit cells), while the total system size was increased along with the number of processors. The results of the scaling analysis are summarized in Fig. 9. As can be seen by comparison with the dashed line, in the weak scaling limit, the acceptance probability scales better than logarithmically with the number of processors. These results provide clear evidence that the VC-SGC-MC algorithm is ideally suited for simulations of very large systems.

The good scalability of the algorithm can be rationalized as follows: in the first part of a VC-SGC-MC trial step, a

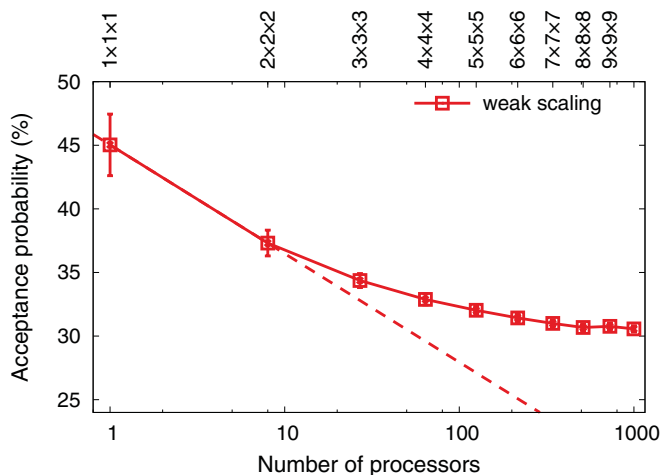


FIG. 9. (Color online) Acceptance probability as a function of the number of processors both in the weak-scaling limit using $\Delta\mu_0 = 0$, $\kappa = 10$, and $c_0 = 0.5$. The dashed line represents logarithmic scaling.

composite move is constructed that in the second part is accepted or rejected as a whole. The collective acceptance/rejection of a large number of individual moves could suggest that the acceptance probability for the second rejection decreases rapidly with the number of individual moves and thus the number of processors. The first acceptance/rejection, however, ensures that the combination of the individual moves form a cluster move that is already “optimized” and therefore has a relatively low probability to be rejected in the second part of the VC-SGC-MC trial move.

IV. APPLICATION TO REALISTIC ALLOYS

A. Sampling structural relaxation and vibrations

In the previous sections, we have discussed in detail the development of an efficient parallel MC algorithm for studying systems with millions of particles at arbitrary global concentrations. For the purpose of demonstration, we considered a simple lattice model. In many practical applications, however, the configuration space includes continuous particle coordinates leading to structural relaxations and thermal vibrations.

As shown in Sec. II A, structural and chemical degrees of freedom can be separated readily in the partition function. This allows us to sample the corresponding integrals with different techniques. A straightforward approach is to combine transmutation and displacement MC trial moves. In practice, this algorithm, however, often converges poorly especially when structural relaxations are involved. As indicated after Eq. (3), a much more efficient way to sample displacements is obtained by combining transmutation Monte Carlo moves with molecular dynamics simulations. In practice, one carries out a MD simulation that is interrupted every n th MD step to execute m MC trial moves. While optimal sampling is obtained if $n = m = 1$ [compare comment after Eq. (3)], for computational efficiency it is beneficial to choose larger values. This does not affect sampling as long as the total number of MD/MC cycles is sufficiently large, i.e., n is much smaller than the total number of MD steps.

We have applied the hybrid MC/MD approach for modeling chemical ordering and/or precipitation in several metallic alloys in the vicinity of heterogeneities such as dislocations, grain boundaries and surfaces. In the next section, we consider

the precipitation of Cu in Fe-rich Fe-Cu nanocrystals as an illustration for the type of problems that can be studied using our algorithm. Other examples include the study of grain boundary pinning in Cu due to Fe impurities,¹¹ structural phase transformations of Cu precipitates in BCC iron,¹² short-range order in Fe-Cr alloys,¹³ and the properties of helium bubbles in Fe and Fe-Cr alloys¹⁴ can be found in Ref. 13, where we used a preliminary version of the present algorithm to study short-range order in Fe-Cr alloys as a function of temperature and composition.

B. Cu precipitation in Fe nanocrystals

We will now concern ourselves with VC-SGC-MC/MD simulations of Cu precipitation in dilute nanocrystalline ferritic Fe-Cu alloys. The very small solubility of Cu in Fe (0.07% at 700 K) gives rise to a very strong driving force for precipitation. The different crystal structures of Fe (body-centered cubic, BCC) and Cu (face-centered cubic, FCC) as well as the mechanical instability of bulk BCC-Cu, imply that as Cu precipitates grow structural phase transformations occur. This realization in conjunction with the technological importance of Fe-Cu alloys has led to a considerable amount of research in this field (see, e.g., Refs. 15 and 16). Here, we compare the precipitation of Cu in dilute nanocrystalline Fe-Cu alloys simulated using two different interatomic potential models.

First, a nanocrystalline BCC-Fe sample with dimensions of 18.8 nm in all Cartesian directions was created as follows. Eleven randomly oriented BCC seeds (average grain diameter 4 nm) were distributed evenly in the simulation cell and nanocrystallites were constructed by filling the Voronoi volumes around each seed. The resulting grain boundary structure was relaxed using conjugate gradient minimization and subsequently equilibrated at a temperature of 700 K for 500 ps using MD simulations. The final sample contained 548 565 atoms.

VC-SGC-MC/MD simulations were performed at 700 K using $\Delta\mu_0 = -0.60$ eV in Eq. (20), $\kappa = 1000$ in Eq. (21), and a target concentration $c_0 = 4\%$. One MC cycle (equivalent to N_{at} MC trial moves where N_{at} is the number of atoms) was carried out per 20 MD steps. The equations of motion were integrated for 1 200 000 MD steps (including 60 000 MC cycles) using a time step of 2.5 fs. Temperature and

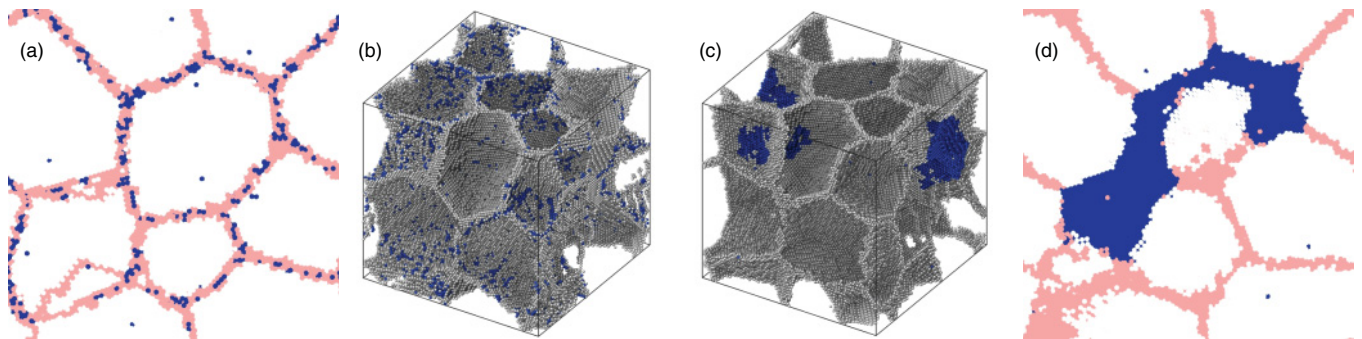


FIG. 10. (Color online) Representative snapshots obtained after full equilibration in simulations using the (a) and (b) LF potential and (c) and (d) the PM potential. Coloring according to common-neighbor analysis. (a) and (d) BCC Fe atoms, Fe and Cu grain-boundary atoms are shown in white, pink, and blue, respectively. (b) and (c) Fe and Cu grain-boundary atoms are shown in gray and blue, respectively.

pressure were maintained using the Nosé-Hoover thermostat and barostat, respectively.

Interatomic interactions were modeled using both the Fe-Cu potential by Ludwig *et al.*¹⁷ (LF) and the potential by Pasianot and Malerba¹⁸ (PM). The LF potential is based on the Fe potential by Simonelli *et al.*¹⁹ and the Cu potential by Voter,²⁰ while the PM potential employs the Fe potential by Mendelev *et al.*²¹ and the Cu potential by Mishin *et al.*²² Both potentials give solubilities at 700 K that are very close to the experimental value (LF: 0.15%, PM: 0.07%, experiment: approximately 0.07%), and thus the target concentration of 4% is far beyond the solubility limits for either potential.

Figure 10 summarizes the key results of our analysis. As expected, both potentials show a very small number of Cu atoms in the center of the grains. As the total Cu concentration of about 4% is far larger than the bulk solubility this implies that excess copper is located in grain boundaries. While the two potentials agree with regard to the latter trend, they yield very different results when it comes to the *distribution* of the Cu in the grain boundaries. Whereas the LF potential predicts a homogeneous distribution with little spatial correlation between the Cu atoms [see Figs. 10(a) and 10(b)], the PM potential yields contiguous Cu precipitates that are agglomerated along only a few neighboring grain boundaries. While this result showcases the kind of insight that can be gained using the VC-SGC-MC/MD hybrid simulation technique, it also demonstrates that further work in the area of potential development and verification is needed before a reliable study of Cu precipitation at grain boundaries in Fe can be conducted.

V. CONCLUSIONS

In the present paper, we have developed a hybrid molecular dynamics/Monte Carlo (MD/MC) algorithm, which is ideally suited for simulating multicomponent systems using samples with millions of particles in both single and multiphase regions of the phase diagram. The most important component is an efficient and scalable transmutation MC method that samples the variance-constrained SGC ensemble. The VC-SGC-MC algorithm can be used to stabilize multiphase equilibria and therefore allows to study precipitation and phase segregation. Since it features a better-than-logarithmic scaling of the acceptance probability with the number of processors, the method is ideally suited for studying very large systems containing several million particles. Finally, by combining the VC-SGC-MC method with molecular dynamics, one obtains a very powerful hybrid scheme that takes into account chemical mixing and precipitation, structural relaxations as well as thermal vibrations.

We have applied the algorithm developed in this work to study the precipitation of Cu in nanocrystalline Fe using two different interatomic potentials. While both potentials predict excess Cu to be located in the grain boundaries, they yield very different results for the distribution of impurity atoms in the grain boundaries. Further work in potential development and verification is required in order to obtain interatomic potential models that provide reliable predictions for element distribution near inhomogeneities such as dislocations, grain boundaries, and surfaces.

The hybrid MC/MC algorithm described in this paper has already been applied to study for example grain boundary pinning in Cu due to Fe impurities,¹¹ structural phase transformations of Cu precipitates in BCC Fe,¹² short-range order in Fe-Cr alloys,¹³ and the properties of helium bubbles in Fe and Fe-Cr alloys.¹⁴ The relation to free energy integration that is implicit to Eq. (A11) has furthermore been utilized in Ref. 9 to obtain the temperature and orientation dependence of free interface energies in Fe-Cr alloys. The algorithms developed in the present work have been implemented in the massively parallel MD code LAMMPS.⁷ The source code is available from the authors.

ACKNOWLEDGMENTS

Lawrence Livermore National Laboratory is operated by Lawrence Livermore National Security, LLC, for the US DOE-NNSA under Contract DE-AC52-07NA27344. Computer time allocations by NERSC at Lawrence Berkeley National Laboratory and the Swedish National Infrastructure for Computing are gratefully acknowledged. One of us (P.E.) has been partly supported by a grant from the Swedish Research Council.

APPENDIX: DERIVATION OF THE VC-SGC ENSEMBLE

In this appendix, we derive the VC-SGC ensemble for the binary system discussed in Sec. II. Consider a system of N particles confined in a box of volume V , where each particle carries a spin of value 0 or 1. Since the VC-SGC ensemble only manipulates the chemical degrees of freedom, we consider for simplicity a system frozen onto a lattice of an arbitrary configuration \mathbf{x}^{3N} . The phase space Ω of this system consists of the set of $\rho = 2^N$ different spin configurations $\{\sigma^N\}$. To simplify the notation below, we enumerate the ρ configurations in Ω : $\{1, 2, \dots, \rho\}$, and thus any spin configuration σ^N can be uniquely identified by its index number.

Let Σ be the set of M representative configurations in Ω and denote by n_α the number of times the α th state appears in Σ . We can uniquely define Σ by the set of numbers $\{n_1, n_2, \dots, n_\rho\}$. The sum of the occupation numbers n_α are constrained according to

$$M = \sum_{\alpha=1}^{\rho} n_\alpha. \quad (\text{A1})$$

We now introduce three more constraints for (i) the average energy \bar{U} , (ii) the average concentration of spin zero particles \bar{c} , and (iii) the square of the concentration of spin zero particles \bar{v}^2 . These constraints can be expressed as

$$\begin{aligned} \bar{U} &= \frac{1}{M} \sum_{\alpha=1}^{\rho} n_\alpha U(\alpha), & \bar{c} &= \frac{1}{M} \sum_{\alpha=1}^{\rho} n_\alpha \hat{c}(\alpha), \\ \bar{v}^2 &= \frac{1}{M} \sum_{\alpha=1}^{\rho} n_\alpha \hat{c}(\alpha)^2. \end{aligned}$$

Above, we have denoted the potential energy for the state α by $U(\alpha)$ and its concentration by $\hat{c}(\alpha)$. For any given set $\Sigma = \{n_\alpha\}$, there are multiple ways of choosing its elements

from Ω . We use this to define the multiplicity η of a set Σ :

$$\eta = \frac{M!}{\prod_{\alpha=1}^{\rho} n_{\alpha}!}.$$

The relative probability of two sets with the same average energy \bar{U} is now determined by the ratio of their multiplicities. In the thermodynamic limit, i.e., large N and large M , the most probable Σ , i.e., the set with the largest multiplicity, will correspond to the equilibrium probability distribution in Ω . Under the above constraints, the most probable distribution of $\{n_{\alpha}\}$ is determined by minimizing the functional $Q(\{n_{\alpha}\}; \mu, \beta, \tilde{\phi}, \tilde{\kappa})$:

$$Q = -\ln \eta - \mu \left(\sum_{\alpha=1}^{\rho} n_{\alpha} - M \right) - \beta \left(\sum_{\alpha=1}^{\rho} n_{\alpha} E(\alpha) - M\bar{U} \right) - \tilde{\phi} \left(\sum_{\alpha=1}^{\rho} n_{\alpha} \hat{c}(\alpha) - M\bar{c} \right) - \tilde{\kappa} \left(\sum_{\alpha=1}^{\rho} n_{\alpha} \hat{c}(\alpha)^2 - M\bar{v}^2 \right). \quad (\text{A2})$$

Above, μ , β , $\tilde{\phi}$, and $\tilde{\kappa}$, are Lagrange multipliers that are introduced as independent variables to facilitate the constrained minimization of the functional Q with respect to the occupation numbers $\{n_{\alpha}\}$. At its minimum, the derivative of the functional Q with respect to the independent variables is set to zero. Setting $\partial Q/\partial n_{\alpha}$ to zero determines their equilibrium distribution:

$$n_{\alpha} = \exp[-\mu - \beta U(\alpha) - \tilde{\phi} \hat{c}(\alpha) - \tilde{\kappa} \hat{c}(\alpha)^2].$$

Using this result in Eq. (A1), we obtain an explicit expression for the chemical potential μ

$$\exp(\mu) = \frac{1}{M} \sum_{\alpha=1}^{\rho} \exp[-\beta U(\alpha) - \tilde{\phi} \hat{c}(\alpha) - \tilde{\kappa} \hat{c}(\alpha)^2]. \quad (\text{A3})$$

Now it is possible to define the equilibrium probability of any state α in Ω as

$$\pi_{\nu}(\alpha) = \mathcal{Z}_{\nu}^{-1} \exp(-\beta\{U(\alpha) + N\hat{c}(\alpha)[\phi + \kappa N\hat{c}(\alpha)]\}), \quad (\text{A4})$$

where $\mathcal{Z}_{\nu} = Me^{\mu}$, and we have introduced the definitions

$$\tilde{\phi} = N\beta\phi, \quad \tilde{\kappa} = N^2\beta\kappa, \quad (\text{A5})$$

in order to reproduce the equilibrium probability distribution of the VC-SGC ensemble Eq. (9). Let us now define the phase space Ω_c of configurations with a fixed concentration c . The canonical free energy $F_C(c)$ for this set can be defined as follows:

$$\exp[-\beta F_C(c)] = \sum_{\alpha \in \Omega_c} \exp[-\beta E(\alpha)]. \quad (\text{A6})$$

In this way the partition function Eq. (A3) can be rewritten as

$$\mathcal{Z}_{\nu} = \int_0^1 \exp\{-\beta[F_C(c) + Nc(\phi + \kappa Nc)]\} dc. \quad (\text{A7})$$

Setting $\partial Q/\partial \tilde{\phi}$ and $\partial Q/\partial \tilde{\kappa}$ in Eq. (A2) to zero and using the definitions (A5) and (A6) provides for a system of two

equations to determine the two unknowns ϕ and κ :

$$\bar{c} = \mathcal{Z}^{-1} \int_0^1 c \exp\{-\beta[F_C(c) + Nc(\phi + \kappa Nc)]\} dc, \quad (\text{A8})$$

$$\bar{v}^2 = \mathcal{Z}^{-1} \int_0^1 c^2 \exp\{-\beta[F_C(c) + Nc(\phi + \kappa Nc)]\} dc. \quad (\text{A9})$$

In solving the above equations, we assume that \bar{v} is chosen such that it is much smaller than \bar{c} and $1 - \bar{c}$. Then it is possible to represent $F_C(c)$ by its Taylor expansion to second order around \bar{c} :

$$F_C(c) = F_C(\bar{c}) + \left. \frac{\partial F_C}{\partial c} \right|_{\bar{c}} (c - \bar{c}) + \frac{1}{2} \left. \frac{\partial^2 F_C}{\partial c^2} \right|_{\bar{c}} (c - \bar{c})^2,$$

and replace the integrals in Eqs. (A7)–(A9) above with indefinite Gaussian integrals

$$1 = \tilde{\mathcal{Z}}_{\nu}^{-1} \int_{-\infty}^{\infty} \exp[-A(c - \bar{c}) - B(c - \bar{c})^2] dc,$$

$$\bar{c} = \tilde{\mathcal{Z}}_{\nu}^{-1} \int_{-\infty}^{\infty} c \exp[-A(c - \bar{c}) - B(c - \bar{c})^2] dc, \quad (\text{A10})$$

$$\bar{v}^2 = \tilde{\mathcal{Z}}_{\nu}^{-1} \int_{-\infty}^{\infty} c^2 \exp[-A(c - \bar{c}) - B(c - \bar{c})^2] dc,$$

where

$$A = \beta \left[\left. \frac{\partial F_C}{\partial c} \right|_{\bar{c}} + N(\phi + 2\kappa N\bar{c}) \right],$$

$$B = \beta \left[\frac{1}{2} \left. \frac{\partial^2 F_C}{\partial c^2} \right|_{\bar{c}} + N^2\kappa \right],$$

$$\tilde{\mathcal{Z}}_{\nu} = \mathcal{Z}_{\nu} \exp\{\beta[F_C(\bar{c}) + N\bar{c}(\phi + \kappa N\bar{c})]\}.$$

It is now easy to see that the system of equations (A10) can be satisfied when

$$A = 0 \quad \text{and} \quad B = \frac{1}{2(\bar{v}^2 - \bar{c}^2)}.$$

Hence within the VC-SGC ensemble, the thermodynamic forces (ϕ and κ) that give rise to a given average concentration \bar{c} and its standard deviation $s_0 = \sqrt{\bar{v}^2 - \bar{c}^2}$, are related to the derivatives of the Helmholtz free energy at \bar{c} as follows:

$$N\phi = \left. \frac{\partial^2 F}{\partial c^2} \right|_{\bar{c}} - \left. \frac{\partial F}{\partial c} \right|_{\bar{c}} - \frac{\bar{c}}{\beta s_0^2},$$

$$N^2\kappa = \frac{1}{2} \left(\frac{1}{\beta s_0^2} - \left. \frac{\partial^2 F}{\partial c^2} \right|_{\bar{c}} \right).$$

The first derivative of the free energy with respect to the concentration of one species, i.e., the difference in chemical potential between the two species $\Delta\mu$, can therefore be calculated from the average concentration according to

$$\Delta\mu \equiv - \frac{1}{N} \left. \frac{\partial F_C}{\partial c} \right|_{\bar{c}} = \phi + 2\kappa N\bar{c}. \quad (\text{A11})$$

We have thus arrived at the important relation Eq. (15), which is used extensively in this paper. In the same way, a similar relation can also be obtained between the second derivative and the variance of the concentration which reads

$$- \left. \frac{\partial^2 F_C}{\partial c^2} \right|_{\bar{c}} = 2N^2\kappa - \frac{1}{\beta s_0^2}. \quad (\text{A12})$$

*sadigh1@llnl.gov

†erhart@chalmers.se

- ¹D. Frenkel and B. Smit, *Understanding Molecular Simulation* (Academic Press, London, 2001).
- ²M. S. Daw and M. I. Baskes, *Phys. Rev. B* **29**, 6443 (1984).
- ³J. Tersoff, *Phys. Rev. Lett.* **56**, 632 (1986).
- ⁴F. H. Stillinger and T. A. Weber, *Phys. Rev. B* **31**, 5262 (1985).
- ⁵J. H. Hetherington, *J. Low Temp. Phys.* **66**, 145 (1987).
- ⁶R. S. Johal, A. Planes, and E. Vives, *Phys. Rev. E* **68**, 056113 (2003).
- ⁷S. Plimpton, *J. Comp. Phys.* **117**, 1 (1995).
- ⁸T. L. Hill, *Thermodynamics of Small Systems* (Dover, 2002).
- ⁹B. Sadigh and P. Erhart, e-print [arXiv:1012.5082](https://arxiv.org/abs/1012.5082) [cond-mat].
- ¹⁰[www.mcs.anl.gov/mpi].
- ¹¹L. A. Zepeda-Ruiz, G. H. Gilmer, B. Sadigh, A. Caro, T. Oppelstrup, and A. V. Hamza, *Appl. Phys. Lett.* **87**, 231904 (2005).
- ¹²P. Erhart, J. Marian, and B. Sadigh (unpublished).
- ¹³P. Erhart, A. Caro, M. Serrano de Caro, and B. Sadigh, *Phys. Rev. B* **77**, 134206 (2008).
- ¹⁴A. Caro, J. Hetherly, A. Stukowski, M. Caro, E. Martinez, S. Srivilliputhur, L. Zepeda-Ruiz, and M. Nastasi, *J. Nucl. Mater.* **418**, 261 (2011).
- ¹⁵P. J. Othen, M. L. Jenkins, G. D. W. Smith, and W. J. Phythian, *Phil. Mag. Lett.* **64**, 383 (1991).
- ¹⁶J. J. Blackstock and G. J. Ackland, *Phil. Mag. A* **81**, 2127 (2001).
- ¹⁷M. Ludwig, D. Farkas, D. Pedraza, and S. Schmauder, *Modell. Simul. Mater. Sci. Eng.* **6**, 19 (1998).
- ¹⁸R. C. Pasianot and L. Malerba, *J. Nucl. Mater.* **360**, 118 (2007).
- ¹⁹G. Simonelli, R. Pasianot, and E. J. Savino, *Mater. Res. Soc. Symp. Proc.* **291**, 567 (1993).
- ²⁰A. F. Voter, Los Alamos National Laboratory Unclassified Technical Report 93-3901, 1993 (unpublished).
- ²¹M. I. Mendeleev, S. Han, D. J. Srolovitz, G. J. Ackland, D. Y. Sun, and M. Asta, *Phil. Mag.* **83**, 3977 (2003).
- ²²Y. Mishin, M. J. Mehl, D. A. Papaconstantopoulos, A. F. Voter, and J. D. Kress, *Phys. Rev. B* **63**, 224106 (2001).