Chemomechanics in alloy phase stability

Sesha Sai Behara [1], John C. Thomas,[1] Brian Puchala,[2] and Anton Van der Ven[1,*]

[1]*Materials Department, University of California, Santa Barbara, California 93106, USA*

[2]*Department of Materials Science and Engineering, University of Michigan, Ann Arbor, Michigan 48109, USA*

We describe a first-principles statistical mechanics method to calculate the free energies of crystalline alloys that depend on temperature, composition, and strain. The approach relies on an extension of the alloy cluster expansion to include an explicit dependence on homogeneous strain in addition to site occupation variables that track the degree of chemical ordering. The method is applied to the Si-Ge binary alloy and is used to calculate free energies that describe phase stability under arbitrary epitaxial constraints. We find that while the incoherent phase diagram (in which coexisting phases are not affected by coherency constraints) hosts a miscibility gap, coherent phase equilibrium predicts ordering and negative enthalpies of mixing. Instead of chemical instability, the chemomechanical free energy exhibits instabilities along directions that couple the composition of the alloy with a volumetric strain order parameter. This has fundamental implications for phase field models of spinodal decomposition as it indicates the importance of gradient energy coefficients that couple gradients in composition with gradients in strain.

I. INTRODUCTION

Strain is increasingly used as an explicit thermodynamic boundary condition to manipulate the properties of materials. Most materials, however, exhibit some chemical complexity as one or more of their sublattices can host multiple chemical species that may order at low temperatures and form disordered solid solutions at high temperatures. There is, therefore, the possibility of important and interesting chemomechanical couplings that can emerge, whereby changes in composition can affect elastic properties and changes in strain can affect chemical interactions.

Early studies by Larche and Cahn [1–3] and Voorhees and Johnson [4] set up a phenomenological framework that rigorously integrates chemical alloy thermodynamics with elasticity and showed that coherency constraints can have a pronounced effect on phase stability and the topology of phase diagrams. The effect of epitaxial constraints on phase stability has been found to be especially strong in semiconductor alloys. Many semiconductor alloys that exhibit miscibility gaps in their equilibrium temperature versus composition phase diagrams can form ordered compounds when epitaxially constrained [5–7]. More generally, the coupling between strain, composition, and degree of ordering in epitaxially grown semiconductor alloys affects phase stability, solubility, and their resulting optoelectronic properties [8–16]. Coherency constraints also affect the electrochemical properties of battery electrode materials as they undergo phase transitions [17] as well as the shapes and compositions of coherent precipitates in metal alloys [18–23] and the electrocatalysis activity and ferroelectric properties of perovskite thin films [24].

First-principles statistical mechanics treatments of chemical order-disorder phenomena usually neglect strain as a thermodynamically controlled boundary condition. The cluster expansion formalism introduced by Sanchez *et al.* [25] is a mathematically rigorous surrogate model that enables an efficient and compact description of the energy of a crystal as a function of the degree of ordering among its different chemical constituents (or of the orientational arrangements of magnetic moments [26–28] or molecules [29]). An alloy cluster expansion that has been trained to first-principles electronic structure calculations can be used to predict alloy thermodynamic properties at finite temperature with mean field or Monte Carlo methods [30,31]. The cluster expansion surrogate model is either used to parametrize the configurational energy of an alloy at constant volume [32] or in the fully relaxed state [33,34] (i.e., at zero pressure) but is not able to generate information about the dependence of thermodynamic properties on strain. Past treatments of the effect of strain on chemical equilibrium in crystalline alloys using the mixed basis cluster expansion of Laks *et al.* [35] account for the long-range interactions that arise from coherency strains, but treat strain implicitly and not as an explicit degree of freedom. The cluster expansion approach has also been extended to describe the dependence of tensor properties such as elastic moduli on the chemical degree of ordering within a crystal [36,37]. While this makes it possible to calculate elastic moduli as a function of temperature and composition, it only provides derivative properties of a full chemomechanical free-energy description of a solid.

In this contribution, we describe a strain plus configuration cluster expansion that has an explicit dependence on both the chemical order and the homogeneous strain of an alloy. The approach enables the calculation of free energies that are functions of temperature, composition, and strain, thereby

*avdv@ucsb.edu

rigorously encapsulating the chemomechanical couplings that affect all thermodynamic properties, including chemical potentials, elastic moduli, and phase stability. The free energies are also well suited to feed into phase field models of microstructure evolution, coherent spinodal decomposition, and precipitation due to coherent ordering reactions [31,38–42]. To illustrate the approach, we calculate the chemomechanical thermodynamic properties of the Si-Ge binary alloy, which can form a disordered solid solution over the sites of the diamond parent crystal structure.

II. METHODS

A. Chemical and strain descriptors

Chemical degrees of freedom in a crystal can be tracked with occupation variables assigned to each crystal site. For a binary A-B alloy, the occupation variables σ_l attached to site l can, for example, take a value of $+1$ if occupied by A and -1 if occupied by B. The chemical arrangement of the whole crystal is then uniquely specified with a vector of occupation variables $\vec{\sigma} = (\sigma_1, \ldots, \sigma_l, \ldots, \sigma_N)$, where N is the total number of sites that can be decorated with A and B atoms.

The degree with which a crystal is homogeneously deformed can be tracked with six independent metrics of strain, $E_{xx}, E_{yy}, E_{zz}, E_{yz}, E_{zx}$, and E_{xy}. These measure the deformation of a crystal relative to its dimensions in a reference state and are defined to be independent of rigid rotations [43]. The numerical values of each strain component depend on the chosen orientation of the Cartesian coordinate system. The optimal orientation is usually guided by symmetry considerations of the reference state used to define strain. For a crystal that has cubic symmetry in the unstrained state, for example, the Cartesian axes are usually chosen to be parallel to the lattice vectors of the cubic unit cell.

There are several definitions of finite strain. The general approach of describing finite strains starts with the deformation gradient, a 3×3 matrix $\mathbf{F}$, that for homogeneous deformations of a crystal relates the lattice vectors of the reference crystal, $\vec{a}, \vec{b}$, and $\vec{c}$, to the lattice vectors of the deformed crystal, $\vec{A}, \vec{B}$, and $\vec{C}$, according to

$$[\vec{A}, \vec{B}, \vec{C}] = \mathbf{F}[\vec{a}, \vec{b}, \vec{c}]. \tag{1}$$

In this relationship, the Cartesian coordinates of the lattice vectors are collected as columns in 3×3 matrices represented by $[\vec{A}, \vec{B}, \vec{C}]$ and $[\vec{a}, \vec{b}, \vec{c}]$. While the deformation gradient $\mathbf{F}$ contains information about the degree to which the crystal has been deformed, it is also affected by any rigid rotation of the deformed crystal. To eliminate the dependence on rigid rotations, metrics of strain are defined in terms of the rotation invariant quantity $\mathbf{F}^{\mathsf{T}}\mathbf{F}$. The Green-Lagrange strain, for example, takes the form $\mathbf{E} = 1/2(\mathbf{F}^{\mathsf{T}}\mathbf{F} - \mathbf{I})$, where $\mathbf{I}$ is the identity matrix, while the Hencky strain is defined as $\mathbf{E} = 1/2 \ln(\mathbf{F}^{\mathsf{T}}\mathbf{F})$.

It is often convenient to work with symmetry-adapted strain order parameters instead of the Cartesian strains directly since order parameters provide more insight into the symmetries that are broken due to a particular strain. When the reference state of the crystal has cubic symmetry, an especially descriptive set of strain order parameters can be defined according to [43]

$$\begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \end{pmatrix} = \begin{pmatrix} \frac{(E_{xx}+E_{yy}+E_{zz})}{\sqrt{3}} \\ \frac{(E_{xx}-E_{yy})}{\sqrt{2}} \\ \frac{(2E_{zz}-E_{xx}-E_{yy})}{\sqrt{6}} \\ \sqrt{2}E_{yz} \\ \sqrt{2}E_{xz} \\ \sqrt{2}E_{xy} \end{pmatrix}. \tag{2}$$

The first strain order parameter e_1 is invariant to symmetry and measures volumetric changes of the crystal as a result of deformation. The next two order parameters, e_2 and e_3, together form an irreducible subspace when the reference crystal has cubic symmetry and describe the tetragonal and orthorhombic distortions of the cubic reference. The last three strain order parameters, e_4, e_5, and e_6, are shear strains. One of the advantages of working with the Hencky strain metric is that the first strain order parameter e_1 is exclusively determined by the relative change in volume of the crystal and any variation of the other strain order parameters, $e_2, \ldots, e_6$, at constant e_1 corresponds to symmetry breaking deformations at constant volume [43].

When measuring the state of strain of an alloy, it is convenient to use a reference at a fixed concentration, such as the dimensions of the unit cell of one of the pure end members. The strain of the crystal can then arise as a result of the imposition of mechanical boundary conditions, such as externally imposed stresses and epitaxial constraints, or due to a change in composition, which may change the equilibrium lattice parameters at fixed pressure.

B. Coupling chemical and strain degrees of freedom

In formulating a strain plus configuration cluster expansion, we consider a binary A-B alloy, but the approach can be generalized to an arbitrary number of chemical constituents [44]. Following the derivation of Sanchez *et al.* [25], the cluster expansion expression for the energy of a crystal can be generalized to explicitly depend on strain $\vec{e}$, in addition to configuration $\vec{\sigma}$, according to

$$E(\vec{e}, \vec{\sigma}) = \sum_{\alpha} \sum_{n_\alpha} V_\alpha^{n_\alpha} \Phi_\alpha^{n_\alpha}(\vec{e}, \vec{\sigma}), \tag{3}$$

where

$$\Phi_\alpha^{n_\alpha}(\vec{e}, \vec{\sigma}) = \lambda_\alpha^{n_\alpha}(\vec{e})\phi_\alpha(\vec{\sigma}) \tag{4}$$

are cluster basis functions. The index α labels clusters of sites in the crystal, such as point, pair, triplet, etc. clusters. The sum extends over all possible 2^N clusters of sites within a crystal, where N is the total number of sites in the crystal that can be occupied by A or B atoms. The cluster basis functions $\Phi_\alpha^{n_\alpha}(\vec{e}, \vec{\sigma})$ are products of functions of strain $\lambda_\alpha^{n_\alpha}(\vec{e})$ with polynomials of the occupation variables defined as

$$\phi_\alpha(\vec{\sigma}) = \prod_{l \in \alpha} \sigma_l. \tag{5}$$

The $\phi_\alpha(\vec{\sigma})$ are the cluster basis functions of the original binary alloy cluster expansion introduced by Sanchez *et al.* [25] and are simply a product of occupation variables belonging to the sites of the cluster α. The strain basis functions $\lambda_\alpha^{n_\alpha}(\vec{e})$ can similarly be represented as polynomials of the six strain variables $\vec{e}$. The n_α is an integer index that tracks distinct polynomials of strain. Finally, $V_\alpha^{n_\alpha}$ are expansion coefficients that embed the specific chemical and elastic properties of the alloyed crystal and are referred to as effective cluster interactions (ECI).

The symmetry of the undecorated and unstrained reference parent crystal imposes constraints on the expansion coefficients $V_\alpha^{n_\alpha}$, and the functional form of the basis functions $\Phi_\alpha^{n_\alpha}(\vec{e}, \vec{\sigma})$. This is because any pair of chemical orderings and/or strain states, $(\vec{e}, \vec{\sigma})$ and $(\vec{e}', \vec{\sigma}')$, that can be mapped onto each other upon the application of a space group symmetry operation $\hat{S}$ of the parent crystal structure must have the same energy. Note that the reference lattice used to measure strain for an alloy is independent of composition and is usually chosen to be the dimensions of the unit cell of one of the end members of the binary alloy.

In generating a suitable basis that satisfies all symmetry constraints, it is useful to rely on several group theoretical concepts related to crystal symmetry and clusters. The first is an orbit of clusters, defined as the collection of all clusters of a particular type that are equivalent by symmetry. By starting with a prototype cluster α, such as the nearest-neighbor pair, all other clusters belonging to its orbit Ω_α can be generated upon the application of the space group operations of the parent crystal to the prototype cluster α. Each distinct cluster type (e.g., point cluster, nearest-neighbor-pair cluster, second-nearest-neighbor-pair cluster, etc.) has a corresponding orbit that collects all symmetrically equivalent clusters of that type. Another useful concept is a cluster group. A prototype cluster α of the cluster orbit Ω_α has a cluster group G_α consisting of a subset of space group operations of the parent crystal that maps the cluster α onto itself.

A set of cluster basis functions that ensure the invariance of the energy of the crystal to the symmetry of the parent crystal structure can be generated as follows. For a prototype cluster α of each cluster orbit Ω_α generate a starting basis of strain polynomials in terms of monomials of each strain variable e_i of the form

$$\eta_\alpha^{\vec{p}}(\vec{e}) = \prod_{i=1}^{6} e_i^{p_i},\tag{6}$$

where $\vec{p} = (p_1, p_2, \ldots, p_6)$ and where each individual p_i is a positive integer ranging from 0 to a maximum polynomial order. In this expression, each $e_i^{p_i}$ is a monomial of the strain variable e_i to the p_i^{th} power. The strain polynomials $\eta_\alpha^{\vec{p}}(\vec{e})$ can be made invariant to the cluster group G_α upon the application of the Reynolds operator, defined as

$$\lambda_\alpha^{n_\alpha}(\vec{e}) = \frac{1}{|G_\alpha|} \sum_{\hat{S} \in G_\alpha} \hat{S}\left[\eta_\alpha^{\vec{p}}(\vec{e})\right],\tag{7}$$

where $|G_\alpha|$ corresponds to the number of symmetry operations in the cluster group G_α and $\hat{S}[\eta_\alpha^{\vec{p}}(\vec{e})]$ represents the application of the symmetry operation $\hat{S}$ to the function $\eta_\alpha^{\vec{p}}$. The application of the Reynolds operator to the starting

basis polynomials generated by enumerating all possible non-negative integers p_i in Eq. (6) produces a set of symmetry invariant polynomials [43]. Some of these polynomials may be duplicates or linear combinations of other symmetry invariant polynomials. After discarding the linearly dependent polynomials, a set of symmetry invariant polynomials $\lambda_\alpha^{n_\alpha}(\vec{e})$ can be collected, each labeled with a distinct integer index n_α. More details about the application of symmetry to strain variables and to polynomials of strain, the Reynolds operator, and the determination of a linearly independent set of symmetry invariant polynomials can be found in Thomas and Van der Ven [43].

Once a set of symmetry invariant cluster basis functions $\Phi_\alpha^{n_\alpha}(\vec{e}, \vec{\sigma})$ with $n_\alpha = 1, \ldots, n_\alpha^{\max}$ for a prototype cluster α have been generated, all other symmetrically equivalent cluster basis functions within the crystal $\Phi_\beta^{n_\alpha}(\vec{e}, \vec{\sigma})$ with $\beta \in \Omega_\alpha$, can be generated upon the application of space group operations to the prototype cluster basis functions $\Phi_\alpha^{n_\alpha}(\vec{e}, \vec{\sigma})$. The resulting symmetrically equivalent basis functions can be collected in an orbit of basis functions labeled $\Omega_\alpha^{n_\alpha}$. The basis functions belonging to the same basis function orbit $\Omega_\alpha^{n_\alpha}$, all share the same expansion coefficient $V_\alpha^{n_\alpha}$ in a symmetry invariant cluster expansion, making it possible to rewrite the strain plus configuration cluster expansion of Eq. (3) as

$$E(\vec{e}, \vec{\sigma}) = \sum_{\Omega_\alpha^{n_\alpha}} V_\alpha^{n_\alpha} \left(\sum_{\beta \in \Omega_\alpha} \Phi_\beta^{n_\alpha}(\vec{e}, \vec{\sigma}) \right).\tag{8}$$

In this form, the outer sum is over distinct orbits of cluster basis functions $\Omega_\alpha^{n_\alpha}$ and the inner sum is over all clusters that are symmetrically equivalent to the prototype cluster α.

Table I lists the first 23 prototype terms of a strain plus configurational cluster expansion for a diamond parent crystal structure. The first term in the expansion, Eq. (8), corresponds to the empty cluster $\alpha = \varnothing$, and has exclusively a strain dependence. There are also other basis functions in the expansion that only depend on occupation variables and are strain independent. Most of the basis functions, though, depend on both occupation variables and a subset of strain variables.

The strain plus configuration cluster expansion of Eq. (8) is for the energy of the whole crystal and is an extensive quantity. This expression can be normalized by the number of unit cells of the crystal N_u upon the introduction of correlation functions defined as

$$\xi_\alpha^{n_\alpha}(\vec{e}, \vec{\sigma}) = \frac{1}{N_u m_{\alpha,n_\alpha}} \sum_{\beta \in \Omega_\alpha} \Phi_\beta^{n_\alpha}(\vec{e}, \vec{\sigma}),\tag{9}$$

where m_{α,n_α} is the multiplicity of the cluster basis function $\Phi_\alpha^{n_\alpha}(\vec{e}, \vec{\sigma})$ per unit cell (i.e., $m_{\alpha,n_\alpha} = |\Omega_\alpha^{n_\alpha}|/N_u$). The energy per unit cell $\epsilon(\vec{e}, \vec{\sigma}) = E(\vec{e}, \vec{\sigma})/N_u$, then becomes

$$\epsilon(\vec{e}, \vec{\sigma}) = \sum_{\Omega_\alpha^{n_\alpha}} V_\alpha^{n_\alpha} m_{\alpha,n_\alpha} \xi_\alpha^{n_\alpha}(\vec{e}, \vec{\sigma}).\tag{10}$$

Expressed in this manner, an approach to parametrize the expansion coefficients to the results of first-principles electronic structure training data becomes evident. In practice, the strain plus configuration cluster expansion must be truncated. The energies of a large number of strain and chemical configurations can be calculated with a first-principles electronic structure method and used to determine the expansion

TABLE I. The first 23 prototype cluster basis functions of a configuration plus strain cluster expansion for the diamond parent crystal structure. The e_1, e_2, e_3, e_4, e_5, and e_6 variables are the strain order parameters defined by Eq. (2). The σ represents the occupation variables that are assigned to each crystal site. These take the values of -1 or $+1$ depending on the occupant of the site (i.e., Si or Ge) for the $Ge_{1-x}Si_x$ alloy. The first eleven basis functions have only a strain dependence. The next five basis functions correspond to point cluster terms. Basis functions Φ_{16} to Φ_{18} are prototype basis cluster functions coupling occupants on the nearest-neighbor-pair cluster with indices i and j, while basis functions Φ_{19} to Φ_{22} couple occupants on the next-nearest-neighbor-pair cluster of the diamond parent crystal. The specific form of the strain-dependent polynomials of each cluster basis function depends on the orientation of the cluster relative to the Cartesian coordinate system.

Cluster function notation	Prototype basis function
Φ_0	1
Φ_1	e_1
Φ_2	e_1^2
Φ_3	$\sqrt{1/2}(e_2^2 + e_3^2)$
Φ_4	$\sqrt{1/3}(e_4^2 + e_5^2 + e_6^2)$
Φ_5	e_1^3
Φ_6	$\sqrt{3/2}(e_1 e_2^2 + e_1 e_3^2)$
Φ_7	$3/2(e_2^2 e_3 - 1/3 e_3^3)$
Φ_8	$(e_1 e_4^2 + e_1 e_5^2 + e_1 e_6^2)$
Φ_9	$\sqrt{3/4}(e_2 e_4^2 - e_2 e_5^2 - \sqrt{1/3} e_3 e_4^2 - \sqrt{1/3} e_3 e_5^2 + \sqrt{4/3} e_3 e_6^2)$
Φ_{10}	$\sqrt{6} e_4 e_5 e_6$
Φ_{11}	σ_i
Φ_{12}	$\sigma_i e_1$
Φ_{13}	$\sigma_i e_1^2$
Φ_{14}	$\sigma_i \sqrt{1/2}(e_2^2 + e_3^2)$
Φ_{15}	$\sigma_i \sqrt{1/3}(e_4^2 + e_5^2 + e_6^2)$
Φ_{16}	$\sigma_i \sigma_j$
Φ_{17}	$\sigma_i \sigma_j e_1$
Φ_{18}	$\sigma_i \sigma_j \sqrt{1/3}(e_4 - e_5 - e_6)$
Φ_{19}	$\sigma_i \sigma_j$
Φ_{20}	$\sigma_i \sigma_j e_1$
Φ_{21}	$\sigma_i \sigma_j \sqrt{3/4}(e_2 - \sqrt{1/3} e_3)$
Φ_{22}	$\sigma_i \sigma_j e_4$

coefficients $V_\alpha^{n_\alpha}$ of a truncated cluster expansion using one of many possible machine-learning regression techniques, such as simple least squares, LASSO, and ridge regression [45–50].

The CASM software package [44,51,52] is capable of algorithmically generating a strain plus configurational cluster expansion for any parent crystal structure. It also has capabilities to enumerate structural models with symmetrically distinct strains and configurations $(\vec{e}, \vec{\sigma})$, and to calculate the correlation functions for each cluster basis function in the configuration $(\vec{e}, \vec{\sigma})$.

C. First-principles electronic structure calculations to parametrize the cluster expansion

The CASM software package was used to enumerate crystallographic models having different arrangements of Si and Ge over the sites of the diamond parent crystal structure and to enumerate different states of strain of each ordered configuration [51]. The formation energy of each enumerated structure was calculated using the Vienna *Ab-initio* Simulation Package (VASP) [53–55]. The density functional theory (DFT) calculations were performed within the generalized gradient approximation (GGA) using the PBE exchange-correlation functional [56]. The projector augmented wave (PAW) method [57] as implemented within VASP with Ge_d and Si pseudopotentials was used. A plane-

wave energy cutoff of 400 eV was used for all the calculations. An automatic Γ-centered k-point mesh was generated for all the structures using a length parameter R_k of 35. A value of 10^{-5} eV was used as an electronic convergence criterion. To achieve geometric convergence, a cutoff value of 0.02 eV/Å was used for forces on all atoms.

III. RESULTS

The strain plus configuration cluster expansion approach was applied to study finite-temperature phase stability and chemical disorder in the Si-Ge alloy. Si and Ge adopt the diamond crystal structure and form high-temperature solid solutions. The lattice parameters of pure Si differ substantially from those of Ge, indicating that coherency strains play an important role in determining phase stability.

A. Formation energies, relaxation strains, and cluster expansion

Figure 1(a) shows the formation energies of 119 fully relaxed Si-Ge configurations on the sites of the diamond parent crystal as a function of the Ge composition. The reference states are pure Si and pure Ge in the diamond crystal structure. The positive formation energies signify a thermodynamic tendency for phase separation and the existence of a miscibility gap in the equilibrium temperature versus composition
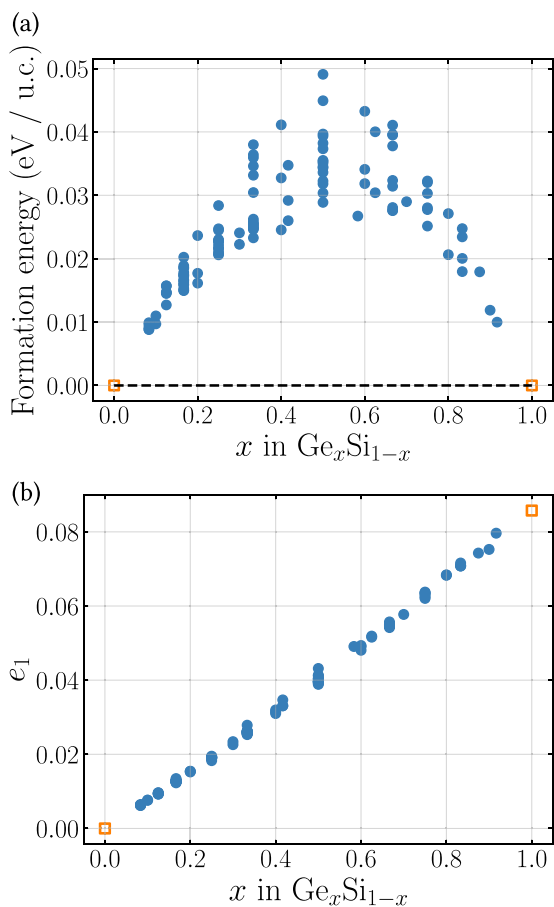
FIG. 1. (a) Formation energies of fully relaxed Si-Ge configurations. Si and Ge in the fully relaxed diamond parent crystal structure were used as reference states. (b) e_1 strain order parameter for fully relaxed Si-Ge configurations with pure Si as the reference state.

phase diagram at constant pressure. Figure 1(b) shows the values of the e_1 strain order parameter for each fully relaxed Si-Ge configuration, calculated using the lattice parameter of pure Si in the diamond crystal as the reference state. The plot shows a very strong and almost linear volume change with increasing Ge concentration, but a very weak dependence of the volume on the nature of the Si-Ge ordering as there is very little variation in e_1 values at each composition.

The strong dependence of the equilibrium volume of the Si-Ge alloy with Ge concentration indicates that coherency constraints should have a significant effect on phase stability. This is manifested in Fig. 2(a), which shows the formation energies of the 119 Si-Ge configurations calculated at a volume that is fixed to the approximate equilibrium volume of the alloy at $x = 0.5$. While the unit-cell dimensions of each configuration were held fixed, all internal atomic coordinates were allowed to relax in these calculations. Figure 2(a) shows that the constraint of a constant volume set at its x = 0.5 value severely penalizes the energies of the Si rich and the Ge rich configurations. Instead of exhibiting a driving force for phase separation, the constant volume formation energies now show
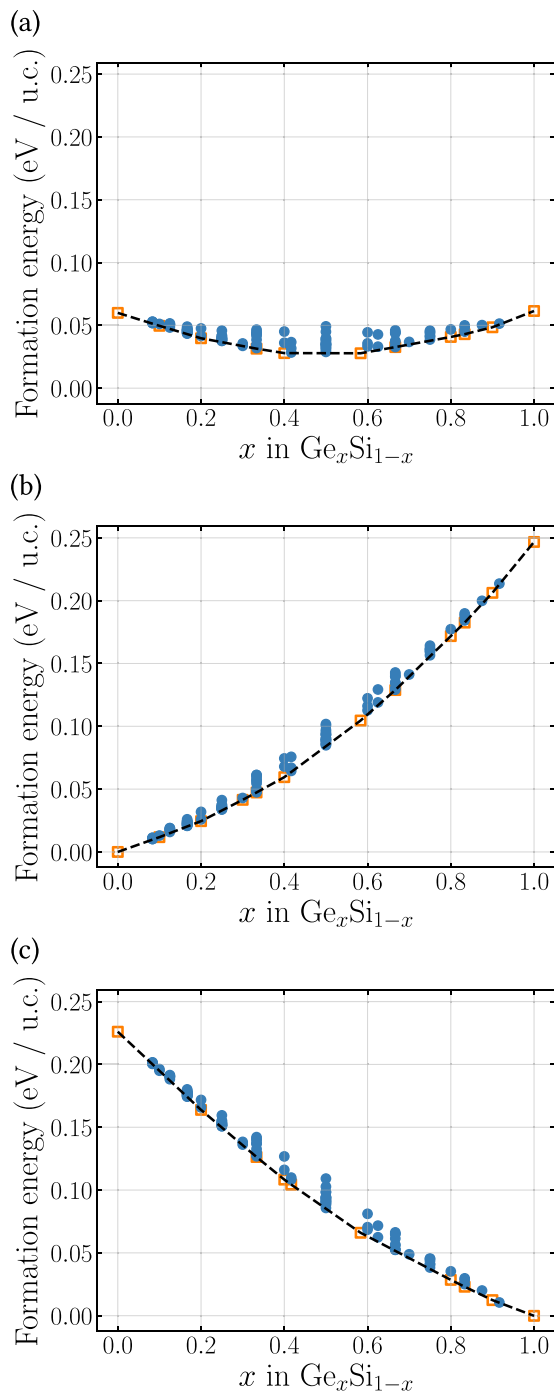

FIG. 2. Formation energies of various Si-Ge configurations at fixed equilibrium volumes of (a) a Si-Ge alloy at $x = 0.5$, (b) pure Si, and (c) pure Ge. The dotted black line represents the thermodynamic convex hull whereas the orange squares represent the ground-state configurations that fall on the convex hull. The formation energies in all three plots are calculated relative to the energies of pure Si and Ge at their equilibrium volumes. Formation energies calculated at fixed volume are always higher than the fully relaxed formation energy. The two formation energies for a given structure are only equal when the imposed strain constraints (including volume) are equal to the fully relaxed strains.

ordering/mixing preferences between Si and Ge. The purely chemical interactions between Si and Ge at constant volume are therefore slightly attractive. Similar mixing formation energies are predicted when fixing the volume to that of pure Si [Fig. 2(b)] or to that of pure Ge [Fig. 2(c)].

The results of Fig. 2 show that there is a strong coupling between chemistry and mechanics in the Si-Ge alloy. To treat this coupling explicitly, we parameterized a strain plus configuration cluster expansion by training to the energies of a large number of systematically strained Si-Ge orderings. For each of the 119 orderings, a grid of symmetrically distinct strain states was enumerated within the e_1 irreducible subspace, within the e_2-e_3 irreducible subspace at each e_1 strain and within the e_4, e_5, and e_6 irreducible subspace at each e_1 and e_2-e_3 strains. The energies of a total of 14 395 chemical plus strain configurations were calculated, allowing only relaxations of internal atomic coordinates. A compact configuration plus strain cluster expansion was then fit to these configurations. The configuration plus strain cluster expansion was able to reproduce the training data with high fidelity, having a root mean squared error (rmse) of 0.0021 eV per unit cell (or 0.001 eV per atom). Only 53 ECI are needed in a strain plus configuration cluster expansion to achieve this accuracy. Figures 3(a) and 3(b) show the rmse of the cluster expansion changes when the next basis function in the expansion is added as part of the truncated cluster expansion. An rmse of 0.0035 eV per unit cell can be achieved with only 23 terms in the cluster expansion (these are terms that extend up to the second-nearest neighbor).

The resultant strain plus configuration cluster expansion was tested on the fully relaxed structures. Each fully relaxed structure has a particular chemical ordering $\vec{\sigma}$, and a relaxation strain $\vec{e}$ relative to an initial reference state. Inserting these values into the strain plus configuration cluster expansion makes it possible to predict the energy of the fully relaxed state. We find that the strain plus configuration cluster expansion is capable of reproducing the fully relaxed energies (to which it was not explicitly trained) with an rmse of 0.0025 eV per unit cell (0.001 eV per atom).

To compare the strain plus configuration cluster expansion to a traditional alloy cluster expansion, we also parameterized a configuration only cluster expansion by training to the fully relaxed energies. A simple least squares method was used to fit a configuration only cluster expansion to the fully relaxed formation energies of 119 Si-Ge orderings [Fig. 1(a)]. A constant, point, and nine pair interactions were sufficient to achieve an rmse of 0.0025 eV per unit cell. Figure 3(c) shows how the rmse of a configuration only cluster expansion changes when basis functions of successively larger clusters are added to the truncated cluster expansion. While a coupled strain and occupational cluster expansion reached an accuracy of 0.0035 eV per unit cell by only including terms that extend up to the second nearest neighbor, the configuration only cluster expansion requires interactions that extend to at least the fifth-nearest neighbor to achieve a similar accuracy. This is because the configuration only cluster expansion has to learn not only the chemical interactions among Si and Ge, but also the implicit dependence of the relaxation strain energy on the arrangement of Si and Ge.
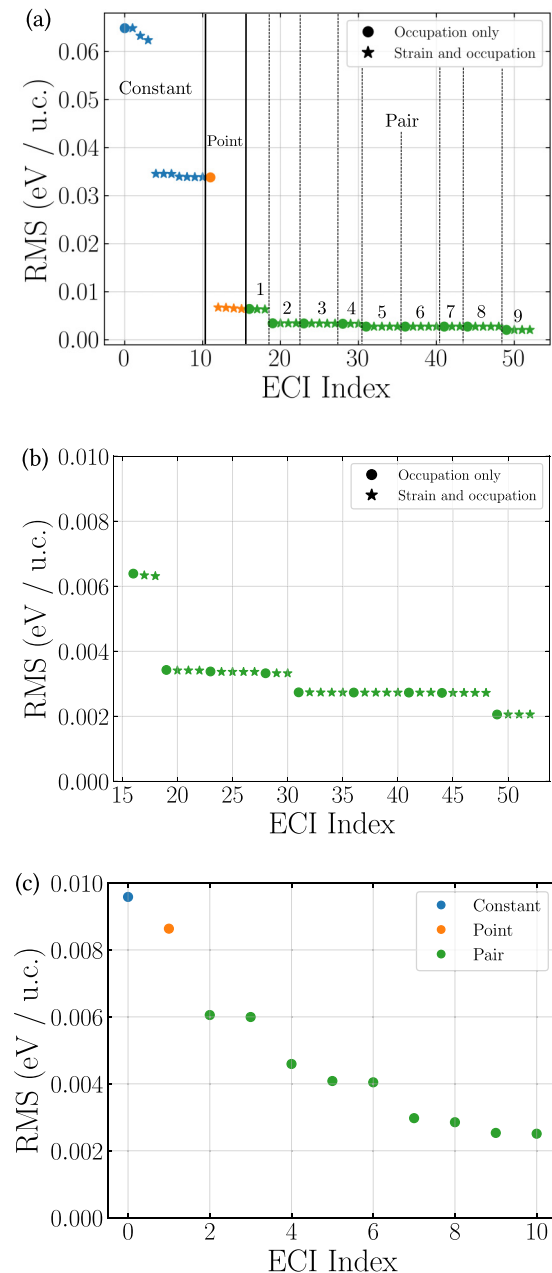

FIG. 3. (a) The root mean square error (rmse) upon the incremental addition of cluster basis functions to a truncated configuration plus strain cluster expansion. (b) Same as (a), but within an rmse range of 0 to 10 meV/unit cell. (c) rmse upon the incremental addition of cluster basis functions for a configuration only cluster expansion.

B. Chemomechanical free energies

The strain plus configuration cluster expansion was used within grand canonical Monte Carlo simulations to calculate the thermodynamic properties of the $Si_{1-x}Ge_x$ alloy as a function of composition, strain, and temperature. By holding the strain $\vec{e}$ in the strain plus configuration cluster expansion constant within the Monte Carlo simulations, it is possible to calculate the Helmholtz free energy as a function of temperature and alloy composition using conventional free-energy
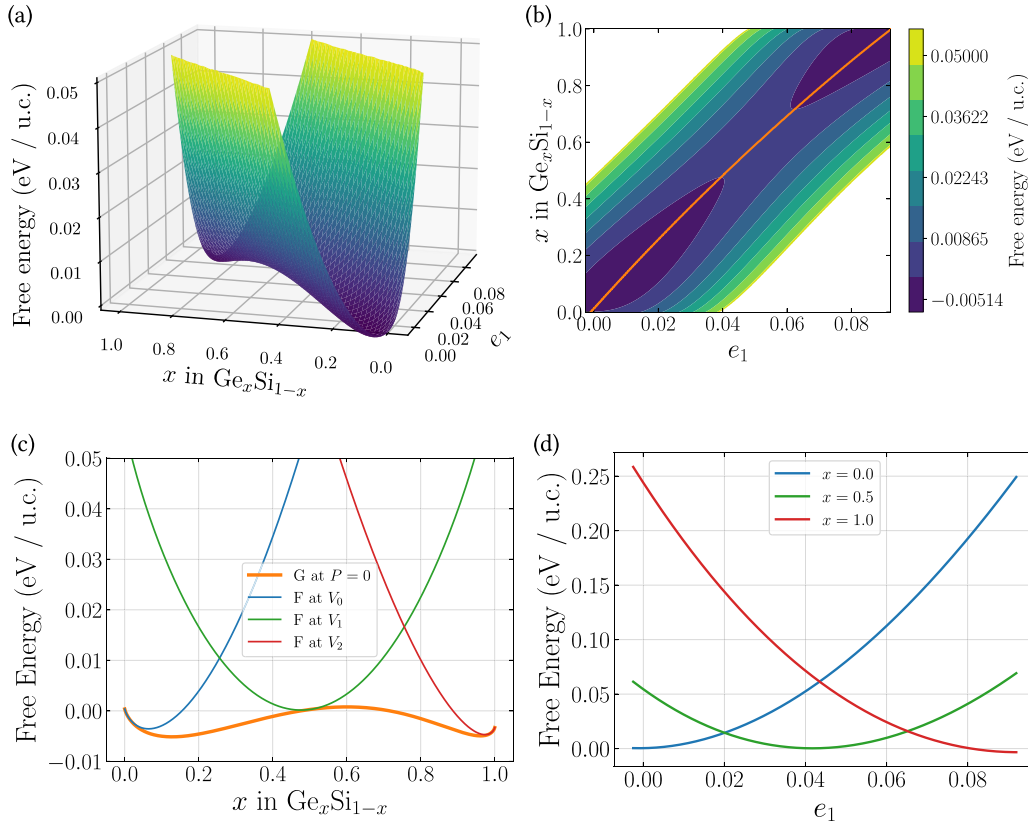
FIG. 4. Helmholtz free energy as a function of concentration x and volumetric strain e_1 calculated at 298 K. (a) A three-dimensional depiction of the free-energy surface and (b) a contour plot. The solid orange line in (b) traces the minimum free-energy path in x-e_1 space, while the minimum free energy at each composition is plotted as a function of concentration in orange in (c). (c) Free-energy slices at different values of e_1 as a function of composition (d) Free-energy slices at different values of composition as a function of e_1.

integration techniques [52,58]. By repeating this over a grid of strains, a Helmholtz free energy per unit cell $f(T, x, \vec{e})$, as a function of temperature T, alloy composition x, and strain $\vec{e}$, can be assembled.

Figure 4(a) shows the calculated Helmholtz free energy at 298 K as a function of alloy composition x and the e_1 strain order parameter, which is a measure of a symmetry preserving volume change. The free-energy surface has two minima, one at Si rich compositions and low values of e_1 and another at Ge rich compositions and high values of e_1. Figure 4(b) shows a contour plot of the free-energy surface as a function of x and e_1. The solid-orange curve in Fig. 4(b) traces out the minimum in the free energy with respect to e_1 for every fixed concentration x, i.e., $e_1(x)$ that satisfies $(\partial f/\partial e_1)_{x,T} = 0$. This is the value of e_1 at which the solid has a zero pressure, as can be seen from the thermodynamic equation of state relationship

$$P = -\left(\frac{\partial F}{\partial V}\right)_{N_{Si},N_{Ge},T} = -\left(\frac{\partial f}{\partial e_1}\right)_{x,T}\left(\frac{\partial e_1}{\partial v}\right), \quad (11)$$

where F is the free energy of the solid, V is the volume, and N_{Si} and N_{Ge} are the number of Si and Ge atoms, respectively. For a solid consisting of N_u unit cells, $F = N_u f$ and $V = N_u v$, where v is the volume per unit cell. The other strain order

parameters $e_2, \ldots, e_6$ are implicitly held constant in all the partial derivatives.

The strain states (e_1) corresponding to the solid orange curve in Fig. 4(b) can be related to the equilibrium lattice parameters of a Si-Ge alloy offering a basis for comparison with experimental observations. The calculated equilibrium lattice parameters at 298 K of pure Si, pure Ge, and an alloy at $x = 0.506$ are 5.4651, 5.7665, and 5.6017 Å respectively. These are in good agreement with experimentally measured values at room temperature by Dismukes *et al.* [59], which report values of 5.431, 5.6575, and 5.5475 Å respectively.

Since atmospheric pressure is very close to zero relative to typical elastic stresses in the solid state, it is common in the *ab initio* literature to approximate equilibrium at atmospheric pressure as equivalent to equilibrium at zero pressure. The Helmholtz free energy along $e_1(x)$ satisfying $(\partial f/\partial e_1)_{x,T} = 0$ then also coincides with the Gibbs free energy of the solid at zero pressure since $G = F + PV$, which simplifies to $G = F$ when $P = 0$. The orange curve in Fig. 4(c) plots this free energy as a function of composition x. It shows the characteristic double-well free-energy curve that produces a miscibility gap in a temperature versus composition phase diagram. However, this free energy is the unconstrained free energy, coinciding with the zero-pressure equilibrium volume. The miscibility gap can only emerge for an incoherent two-phase mixture,
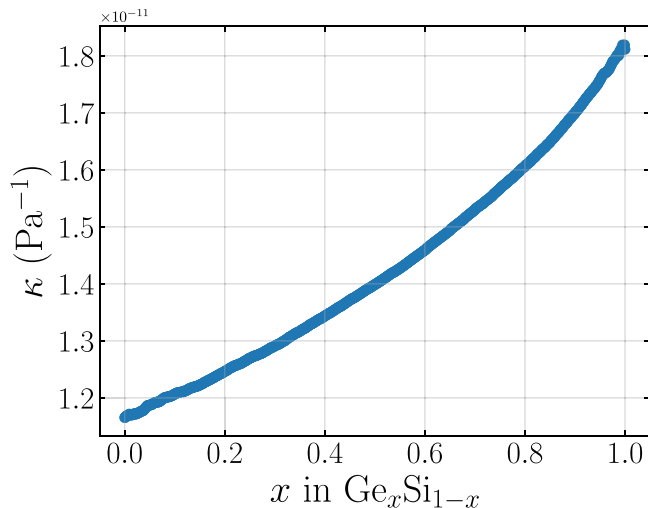
FIG. 5. The calculated compressibility of the Ge_xSi_{1-x} alloy as a function of concentration at $P = 0$.

where the coexisting phases are free to relax to their equilibrium volumes at zero pressure. Also shown in Fig. 4(c) are three free-energy curves as a function of composition, but taken at constant slices of e_1. These free energies have positive curvatures and therefore exhibit a tendency for chemical mixing (as opposed to phase separating). Any attempt to realize a coherent phase separation when the solid is constrained to a fixed value of e_1 results in an increase in the free energy. In the presence of coherency constraints, the Si-Ge alloy will therefore form a disordered solid solution. Each of these curves is tangent to the zero-pressure free-energy curve at the composition x coinciding with the imposed e_1. Figure 4(d) shows free-energy slices at three fixed concentrations plotted as a function of e_1. The curvature of the free energy with respect to e_1 at fixed concentration is related to the compressibility of the solid

$$\kappa = -\frac{1}{v}\left(\frac{\partial v}{\partial P}\right) = \frac{1}{v}\frac{1}{\left(\frac{\partial^2 f}{\partial v^2}\right)} \tag{12}$$

where T, x, N_u (number of unit cells) and the strains, $e_2, \ldots, e_6$ are held constant in the above derivatives. The second derivative of the free energy f with respect to volume can be expressed in terms of derivatives with respect to e_1 according to

$$\left(\frac{\partial^2 f}{\partial v^2}\right) = \left[\left(\frac{\partial^2 f}{\partial e_1^2}\right)\left(\frac{\partial e_1}{\partial v}\right)^2 + \left(\frac{\partial f}{\partial e_1}\right)\left(\frac{\partial^2 e_1}{\partial v^2}\right)\right], \tag{13}$$

where T, x, and the strains, $e_2, \ldots, e_6$ are again held constant. Figure 5 shows the dependence of the compressibility of Ge_xSi_{1-x} as a function of alloy concentration at 298 K. The compressibility of the alloy increases almost linearly between the compressibilities of Si and Ge. The calculated compressibility of pure Si of 1.16×10^{-11} Pa^{-1} is in good agreement with the experimentally measured compressibility of 1.018×10^{-11} Pa^{-1} [60]. The agreement between the calculated compressibility for pure Ge of 1.81×10^{-11} Pa^{-1} and the experimentally reported value of 1.33×10^{-11} Pa^{-1} [61] is not as good as for pure Si. Previous *ab initio* studies

using GGA PBE report similar values for the Ge compressibility (1.78×10^{-11} Pa^{-1}) [62]. Note that properties such as the compressibility are very sensitive to the exchange correlation functional approximation used as part of the DFT calculations. For example, calculations within the local density approximation (LDA) predict compressibilities for pure Si and Ge that are more accurate than those predicted with PBE [62].

The full tensor of elastic moduli of the Ge_xSi_{1-x} alloy are related in a similar way to second derivatives of the chemomechanical free energy $f(T, x, \vec{e})$ with respect to the six independent Cartesian strains E_{xx}, E_{yy}, E_{zz}, E_{xy}, E_{xz}, and E_{yz}. The mathematical relationship between the elastic moduli and the second derivatives of the chemomechanical free energies depends on the type of strain that is used (i.e., Green-Lagrange or a Hencky strain). A simple example where chemomechanical free energies are important is in the context of coherent epitaxial growth of alloys on a "semi-infinite" substrate [Fig. 6(a)]. As a concrete example, we consider a Ge_xSi_{1-x} alloy [green in Fig. 6(a)] grown on the (001) surface of a $Ge_{0.5}Si_{0.5}$ substrate [yellow in Fig. 6(a)]. If the epitaxially grown alloy has a composition that differs from that of the substrate, it will be subjected to epitaxial strains that alter its free energy. The semi-infinite substrate constrains the dimensions of the epitaxial alloy in the $\hat{x} - \hat{y}$ plane, which is parallel to the substrate surface. The substrate therefore fixes the E_{xx}, E_{yy}, and E_{xy} strains of the epitaxial alloy to the equilibrium values of the $Ge_{0.5}Si_{0.5}$ substrate alloy. Perpendicular to the substrate, however, the epitaxial alloy is free to relax along the $\hat{z}$ direction until mechanical equilibrium is reached with the externally imposed stress, which is usually determined by the pressure P of the gaseous environment. For the simple geometry of this example, the stress within the epitaxial alloy along the $\hat{z}$ direction is related to the first derivative of the free energy with respect to the Hencky strain E_{zz} according to

$$\sigma_{zz} = \frac{1}{v}\frac{\partial f}{\partial E_{zz}} = \frac{1}{v}\left[\frac{\partial f}{\partial e_1}\frac{\partial e_1}{\partial E_{zz}} + \frac{\partial f}{\partial e_3}\frac{\partial e_3}{\partial E_{zz}}\right]. \tag{14}$$

In this expression, $v = V/N_u$ is the volume per unit cell in the current state. All derivatives are taken while holding T, N_{Si}, N_{Ge}, and the strains E_{xx}, E_{yy}, E_{yz}, E_{xz}, and E_{xy} constant.

Figure 6(b) shows the calculated chemomechanical free energy as a function of alloy concentration and E_{zz} at the constant strains E_{xx} and E_{yy} imposed by the substrate ($E_{yz} = E_{xz} = E_{xy} = 0$ due to the symmetry of the problem). Figure 6(c) shows slices of the free energy at three different alloy compositions as a function of E_{zz}. When the film is in equilibrium with a low-pressure environment, such as vacuum or atmospheric pressure, $P \approx 0$, and the equilibrium value of E_{zz} can be determined by setting Eq. (14) to zero. This is equivalent to minimizing the chemomechanical free energy of Fig. 6(b) with respect to E_{zz} at each alloy composition. The resulting minimum free energy is shown as the orange curve in Fig. 6(d). This free energy is compared to the free energy of the alloy at $P = 0$, without epitaxial constraints (blue curve). The two curves touch each other at $x = 0.5$ since the epitaxial alloy then has the same composition as the substrate and there are no epitaxial strains. When the composition of the
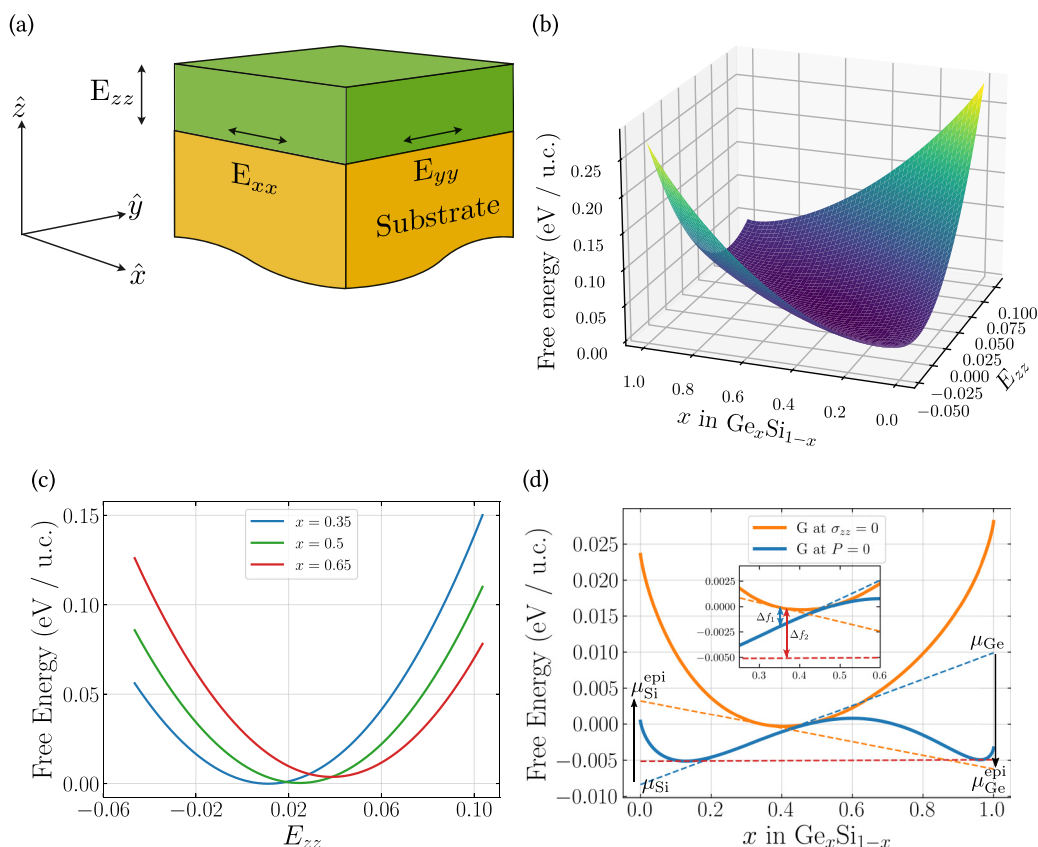
FIG. 6. (a) Schematic illustration of epitaxial growth. (b) Helmholtz free energy at room temperature calculated as a function of composition and E_{zz} under epitaxial strain conditions (fixed E_{xx} and E_{yy}). (c) Free-energy slices at $x = 0.35, 0.5, 0.65$ as a function of E_{zz}. (d) The solid-orange line shows the epitaxially constrained free energy coinciding with $\sigma_{zz} = 0$. The solid-blue line corresponds to the equilibrium free energy when $P = 0$. The dotted-orange line shows the tangent to the epitaxially constrained free energy at $x = 0.35$. The black arrows indicate how the chemical potentials of Si and Ge change upon going from an unconstrained state to an epitaxially strained state at fixed E_{xx} and E_{yy} and $\sigma_{zz} = 0$.

epitaxial alloy differs from that of the substrate, the epitaxial misfit strains in the $\hat{x}$ and $\hat{y}$ directions result in a strain energy cost, which raises the free energy of the epitaxial alloy relative to an unconstrained alloy at the same composition. Hence the free energy of the epitaxial alloy (orange curve) is always above that of the unconstrained alloy at $P = 0$ (blue curve).

The chemical potentials of Si and Ge can be determined graphically from the intercept of the tangent to the free-energy curves with the energy axis at $x = 0$ and $x = 1$, respectively, as shown in Fig. 6(d). The red-dashed line in Fig. 6(d) corresponds to the common tangent to the free-energy curve at zero pressure, which predicts an incoherent miscibility gap at zero pressure. The orange-dashed line is tangent to the free energy of the epitaxially strained alloy at $x = 0.35$. Its intercepts with the free-energy axis at $x = 0$ and $x = 1$ correspond to the Si and Ge chemical potentials in the epitaxially strained alloy. The blue-dashed line is tangent to the free energy of the unconstrained alloy at $x = 0.35$. The Si and Ge chemical potentials of the epitaxially constrained alloy at $x = 0.35$ differ from their values in the unconstrained alloy at the same composition $x = 0.35$. Since the epitaxial alloy at $x = 0.35$ must undergo a positive strain in the $\hat{x}$

and $\hat{y}$ directions ($E_{xx} > 0$ and $E_{yy} > 0$) as it is coherently attached to a substrate having a composition $x = 0.5$ (which has a larger volume), the chemical potential of the smaller Si atoms increases, while the chemical potential of the larger Ge atoms decreases relative to their values in the unconstrained alloy at the same composition. The epitaxial strain energy, therefore, sets up diffusional driving forces that favor the diffusion of Ge (Si) to (away from) the epitaxially strained alloy.

The difference in the free energy between the epitaxial alloy and the unconstrained alloy at $x = 0.35$, denoted Δf_1 in Fig. 6(d), can be interpreted as the strain free energy per unit cell that is needed to elastically deform the unconstrained alloy to achieve epitaxy with the substrate. The total strain free energy of the epitaxial alloy scales with the size of the film and therefore will be linearly proportional to the film thickness. Above a critical thickness, the strain energy will exceed the cost of introducing misfit dislocations along the substrate/film interface that can relieve the epitaxial constraints and allow the alloy to lower its free energy. If the constituents of the epitaxially strained alloy phase separate to form an incoherent two-phase mixture, the free energy will lower by Δf_2 as illustrated in Fig. 6(d).

IV. DISCUSSION

We have introduced a strain plus configuration cluster expansion formalism that enables a rigorous treatment of chemical and strain degrees of freedom in multi-component crystals. The approach makes it possible to calculate free energies that explicitly depend on temperature, composition, and strain. The first derivatives of such free-energy surfaces determine chemical potentials and stresses, while their curvatures with respect to strain variables are related to elastic moduli [63]. A free-energy description that explicitly couples composition with strain, therefore, enables a straightforward calculation of the dependence of the elastic moduli on concentration and short-range order.

The strain variables that appear in the chemomechanical free energy are measured relative to a fixed reference state, independent of concentration. For the Ge_xSi_{1-x} alloy treated in this paper, for example, the reference state for strain was chosen to be the fully relaxed diamond unit cell of pure Si. Since the equilibrium volume and shape of the unit cell can change substantially upon varying the alloy composition, it is important to use measures of finite strain that are invariant to rigid rotations. The Green-Lagrange strain can serve this purpose, but for first-principles free-energy surfaces expressed in terms of symmetry adapted strain order parameters, it is often more convenient to work with the Hencky strain [43,64,65]. This is because the symmetry invariant strain order parameter e_1 is then directly proportional to the volume change, and any variations in the symmetry breaking strain order parameters, $e_2, \ldots, e_6$, at constant e_1, occur at constant volume [43]. Due to the reliance on the same reference state for strain at all compositions, the strain variables can be separated into a chemical strain $e_i^{\text{chem}}(x)$, which depends on the alloy composition, and an elastic strain e_i^{elast}, according to

$$e_i = e_i^{\text{chem}}(x) + e_i^{\text{elast}}. \tag{15}$$

The chemical strain corresponds to the strain that minimizes the free energy at each composition. The elastic strain then measures deformations of the crystal relative to its equilibrium dimensions at that composition given by $e_i^{\text{chem}}(x)$. It is important to note that the strains are all measured relative to a high-symmetry reference state that is independent of composition and usually chosen to be one of the pure end members of the alloy.

The strain plus configuration cluster expansion can be used to treat an alloy whose parent crystal structure is susceptible to a group/subgroup symmetry breaking of its unit cell due to a variation in composition. A common example occurs when a cubic phase undergoes a tetragonal distortion due to an electronic instability induced by a change in the composition of the alloy. The reference state for strain should then be chosen to have the highest symmetry such that all symmetrically equivalent lower symmetry variants can be described with a common strain plus configuration cluster expansion. Many alloys host multiple parent crystal structures [66] that are not related through a group/subgroup symmetry relationship with respect to strain and small atomic displacement degrees of freedom [67]. Treating these systems requires the parametrization of a separate strain plus

configuration cluster expansion for each distinct parent crystal structure.

While contributions to the free energy due to vibrational excitations have been neglected in our study of the Si-Ge alloy, these can be included in a natural way using standard coarse-graining techniques [68]. For each enumerated chemical configuration and strain state, it is possible to calculate the harmonic vibrational free energy [69–72]. The resulting harmonic free-energy surfaces can then be used to train a strain plus configuration cluster expansion with temperature-dependent ECI. Monte Carlo simulations applied to the strain plus configuration cluster expansion will then generate thermodynamic properties that account for configurational and harmonic vibrational excitations. The resultant free-energy surfaces are at the level of a generalized quasiharmonic approximation but depend on the full strain tensor and account for configurational degrees of freedom.

Free-energy surfaces that explicitly depend on composition and strain are the natural ingredients to phase-field simulations [38] of microstructure evolution that couples chemical redistribution through diffusion with elastic driving forces due to coherency strains [39]. A phase field model that describes coherent microstructure evolution departs from a free energy of the form

$$\int_V [f(T, x, \vec{e}) + \gamma(T, x, \vec{e})] \frac{d\vec{r}}{v}, \tag{16}$$

where the first term of the integrand is the homogeneous free energy and the second term collects gradient energy contributions that correct the homogeneous free energy in the presence of gradients in composition and strain. The gradient energy contributions take the form

$$\gamma(T, x, \vec{e}) = \sum_{\alpha,\beta} K_{\alpha,\beta}(x, \vec{e}) \frac{\partial x}{\partial r_\alpha} \frac{\partial x}{\partial r_\beta}$$
$$+ \sum_{i,j} \sum_{\alpha,\beta} \Gamma_{\alpha,\beta}^{i,j}(x, \vec{e}) \frac{\partial e_i}{\partial r_\alpha} \frac{\partial e_j}{\partial r_\beta}$$
$$+ \sum_i \sum_{\alpha,\beta} \Lambda_{\alpha,\beta}^i(x, \vec{e}) \frac{\partial x}{\partial r_\alpha} \frac{\partial e_i}{\partial r_\beta}, \tag{17}$$

where r_α and r_β represent components of a Cartesian vector, and $K_{\alpha,\beta}$, $\Gamma_{\alpha,\beta}^{i,j}$, and $\Lambda_{\alpha,\beta}^i$ are gradient energy coefficients. The symmetry of the parent crystal imposes constraints between the different gradient energy coefficients [73].

Gradient energy contributions are only strictly necessary if the homogeneous free energy exhibits negative curvatures along particular directions in the variable space spanned by x and $e_1, \ldots, e_6$. The gradient energy term then regularizes the homogeneous free energy in phase-field simulations of spinodal decomposition or ordering reactions [74–76]. The calculated chemomechanical free-energy surfaces of the Ge_xSi_{1-x} alloy have positive curvatures with respect to both composition [Fig. 4(d)] and strain [Fig. 4(c)]. It is only along directions that mix composition and the e_1 strain that negative curvatures emerge in the chemomechanical free energy of Ge_xSi_{1-x}, as is evident in Fig. 4(b). A description of spinodal decomposition in the Ge_xSi_{1-x} alloy would then require symmetry allowed gradient energy coefficients $\Lambda_{\alpha,\beta}^i$ that couple gradients in composition with gradients in strain,

i.e., $(\partial x/\partial r_\alpha)(\partial e_i/\partial r_\beta)$. This is an overlooked subtlety in treatments of spinodal decomposition in the presence of coherency strain, where a negative curvature with respect to the composition axis only is generally assumed and gradient energy terms that couple gradients in composition with gradients in strain are neglected [74,75,77]. Some multi-component solids can also exhibit negative curvatures with respect to symmetry breaking strains, allowing for the occurrence of chemomechanical spinodal decomposition [39,63,78]. In this case, strain gradient energy contributions must be included in the free-energy description [79]. The chemomechanical free energy can also be extended to include a dependence on order parameters that track chemical orderings over the parent crystal structure [31,40].

Cluster expansions that only depend on chemical configurational degrees of freedom are unable to accurately treat the long-range effects of coherency strains when the expansion is truncated, instead being most suited for the calculation of the thermodynamic properties of homogeneous phases. The mixed basis cluster expansion approach of Laks *et al.* [35] was designed to rectify this deficiency. While a powerful method to model alloy thermodynamics [80–85] and coherent multi-phase equilibrium at the meso scale [86], the method does not enable the calculation of free energies that explicitly depend on strain. Furthermore, it is considerably more complex than a real-space cluster expansion that can be truncated beyond a maximal cluster size, as it requires the calculation of the constituent strain energy and frequent evaluations of Fourier transforms within Monte Carlo simulations. Finally, while such a treatment can be expected to be rigorous for systems that exhibit simple miscibility gaps, it becomes more questionable when complex ordered phases that cannot be described as simple layered orderings are stable. The strain plus configuration cluster expansion introduced in this paper addresses many of the shortcomings of both the real-space and mixed-basis cluster expansions for solids that can undergo large variations in volume and group/subgroup symmetry breaking strains with composition. As a surrogate model for first-principles statistical mechanics simulations, the strain plus configuration cluster expansion should be a convenient compromise between the more restrictive configuration only cluster expansions [25,30,35] and the more versatile machine-learned interatomic potentials [87–89], which treat all degrees of freedom, but require a substantially larger effort to train with first-principles electronic structure calculations.

V. CONCLUSIONS

We have described a statistical mechanics approach to calculate chemomechanical free energies that depend on temperature, composition, and strain. This is enabled with an extension of the alloy cluster expansion to also include a dependence on homogeneous strain variables. The symmetry of the parent crystal structure imposes constraints on the form of the expansion basis functions. To illustrate the approach, we calculated the chemomechanical free energy of the Si-Ge alloy. In the absence of coherency constraints at zero pressure, the alloy has phase separating tendencies. Coherency constraints, however, modify phase stability and lead to mixing tendencies among Si and Ge on the diamond parent crystal structure. Chemomechanical free-energy surfaces that depend not only on temperature and composition but also explicitly on finite measures of strain are essential ingredients to rigorous phase field models of coherent microstructure evolution.

[1] F. Larché and J. W. Cahn, Acta Metall. **26**, 53 (1978).

[2] F. Larche and J. Cahn, Acta Metall. **26**, 1579 (1978).

[3] F. Larché and J. W. Cahn, Acta Metall. **33**, 331 (1985).

[4] P. W. Voorhees and W. C. Johnson, in *Solid State Physics-Advances in Research and Applications* (Academic Press, New York, 2004), pp. 1–201

[5] A. Zunger and D. Wood, J. Cryst. Growth **98**, 1 (1989).

[6] J. L. Martins and A. Zunger, Phys. Rev. Lett. **56**, 1400 (1986).

[7] D. M. Wood and A. Zunger, Phys. Rev. B **40**, 4062 (1989).

[8] W. Donner, C. Chen, M. Liu, A. J. Jacobson, Y.-L. Lee, M. Gadre, and D. Morgan, Chem. Mater. **23**, 984 (2011).

[9] T. F. Kuech, L. J. Mawst, and A. S. Brown, Annu. Rev. Chem. Biomol. Eng. **4**, 187 (2013).

[10] T. F. Kuech, S. E. Babcock, and L. Mawst, Appl. Phys. Rev. **3**, 040801 (2016).

[11] A. M. Holder, S. Siol, P. F. Ndione, H. Peng, A. M. Deml, B. E. Matthews, L. T. Schelhas, M. F. Toney, R. G. Gordon, W. Tumas *et al.*, Sci. Adv. **3**, e1700270 (2017).

[12] H. Jacobsen, B. Puchala, T. F. Kuech, and D. Morgan, Phys. Rev. B **86**, 085207 (2012).

[13] M. A. Stevens, K. A. Grossklaus, and T. E. Vandervelde, J. Cryst. Growth **527**, 125216 (2019).

[14] Y. Guan, G. Luo, D. Morgan, S. E. Babcock, and T. F. Kuech, J. Phys. Chem. Solids **138**, 109245 (2020).

[15] R. R. Schnepf, J. J. Cordell, M. B. Tellekamp, C. L. Melamed, A. L. Greenaway, A. Mis, G. L. Brennecka, S. Christensen, G. J. Tucker, E. S. Toberer *et al.*, ACS Energy Lett. **5**, 2027 (2020).

[16] J. J. Cordell, M. K. Miller, M. B. Tellekamp, A. Tamboli, G. J. Tucker, and S. Lany, Phys. Rev. Appl. **18**, 064030 (2022).

[17] A. Van der Ven, K. Garikipati, S. Kim, and M. Wagemaker, J. Electrochem. Soc. **156**, A949 (2009).

[18] J. Y. Huh, J. Howe, and W. Johnson, Acta Metall. Mater. **41**, 2577 (1993).

[19] M. Doi, Prog. Mater. Sci. **40**, 79 (1996).

[20] A. R. Natarajan, E. L. Solomon, B. Puchala, E. A. Marquis, and A. Van der Ven, Acta Mater. **108**, 367 (2016).

[21] A. R. Natarajan and A. Van der Ven, Acta Mater. **124**, 620 (2017).

[22] S. DeWitt, E. L. Solomon, A. R. Natarajan, V. Araullo-Peters, S. Rudraraju, L. K. Aagesen, B. Puchala, E. A. Marquis, A. Van Der Ven, K. Thornton *et al.*, Acta Mater. **136**, 378 (2017).

[23] E. L. Solomon, A. R. Natarajan, A. M. Roy, V. Sundararaghavan, A. Van der Ven, and E. A. Marquis, Acta Mater. **166**, 148 (2019).

[24] J. Hwang, Z. Feng, N. Charles, X. R. Wang, D. Lee, K. A. Stoerzinger, S. Muy, R. R. Rao, D. Lee, R. Jacobs *et al.*, Mater. Today **31**, 100 (2019).

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

[26] R. Drautz and M. Fähnle, Phys. Rev. B **69**, 104404 (2004).

[27] E. Decolvenaere, E. Levin, R. Seshadri, and A. Van der Ven, Phys. Rev. Mater. **3**, 104411 (2019).

[28] D. A. Kitchaev, E. C. Schueller, and A. Van der Ven, Phys. Rev. B **101**, 054409 (2020).

[29] J. C. Thomas, J. S. Bechtel, and A. Van der Ven, Phys. Rev. B **98**, 094105 (2018).

[30] D. De Fontaine, in *Solid State Physics*, Vol. 47 (Elsevier, Amsterdam, 1994), pp. 33–176.

[31] A. Van der Ven, J. C. Thomas, B. Puchala, and A. R. Natarajan, Annu. Rev. Mater. Res. **48**, 27 (2018).

[32] J. Sanchez, J. Phase Equilib. Diffus. **38**, 238 (2017).

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

[34] A. Zunger, in *Statics and Dynamics of Alloy Phase Transformations* (Springer, New York, 1994), pp. 361–419.

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

[36] A. van de Walle, Nat. Mater. **7**, 455 (2008).

[37] J. Z. Liu, A. van de Walle, G. Ghosh, and M. Asta, Phys. Rev. B **72**, 144109 (2005).

[38] L.-Q. Chen, Annu. Rev. Mater. Res. **32**, 113 (2002).

[39] S. Rudraraju, A. Van der Ven, and K. Garikipati, npj Comput. Mater. **2**, 16012 (2016).

[40] A. R. Natarajan, J. C. Thomas, B. Puchala, and A. Van der Ven, Phys. Rev. B **96**, 134204 (2017).

[41] G. H. Teichert, A. R. Natarajan, A. Van der Ven, and K. Garikipati, Comput. Methods Appl. Mech. Eng. **353**, 201 (2019).

[42] G. H. Teichert, A. R. Natarajan, A. Van der Ven, and K. Garikipati, Comput. Methods Appl. Mech. Eng. **371**, 113281 (2020).

[43] J. C. Thomas and A. Van der Ven, J. Mech. Phys. Solids **107**, 76 (2017).

[44] J. C. Thomas, B. Puchala, and A. Van der Ven (unpublished).

[45] C. M. Bishop, *Pattern Recognition and Machine Learning*, 1st ed. (Springer Science+Business Media, New York, 2006).

[46] G. L. Hart, V. Blum, M. J. Walorski, and A. Zunger, Nat. Mater. **4**, 391 (2005).

[47] T. Mueller and G. Ceder, Phys. Rev. B **80**, 024103 (2009).

[48] L. J. Nelson, V. Ozoliņš, C. S. Reese, F. Zhou, and G. L. W. Hart, Phys. Rev. B **88**, 155105 (2013).

[49] L. J. Nelson, G. L. W. Hart, F. Zhou, V. Ozoliņš, Phys. Rev. B **87**, 035125 (2013).

[50] D. E. Ober and A. Van der Ven, arXiv:2309.12255.

[51] B. Puchala, J. C. Thomas, A. R. Natarajan, J. G. Goiri, S. S. Behara, J. L. Kaufman, and A. Van der Ven, Comput. Mater. Sci. **217**, 111897 (2023).

[52] B. Puchala, J. C. Thomas, and A. Van der Ven, arXiv:2309.11761.

[53] G. Kresse and J. Hafner, Phys. Rev. B **47**, 558 (1993).

[54] G. Kresse and J. Furthmüller, Comput. Mater. Sci. **6**, 15 (1996).

[55] G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996).

[56] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[57] P. E. Blöchl, Phys. Rev. B **50**, 17953 (1994).

[58] A. van de Walle and M. Asta, Modell. Simul. Mater. Sci. Eng. **10**, 521 (2002).

[59] J. Dismukes, L. Ekstrom, and R. Paff, J. Phys. Chem. **68**, 3021 (1964).

[60] G. Bartl, C. Elster, J. Martin, R. Schödel, M. Voigt, and A. Walkov, Meas. Sci. Technol. **31**, 065013 (2020).

[61] H. McSkimin, J. Acoust. Soc. Am. **30**, 314 (1958).

[62] S. Wang and H. Ye, J. Phys.: Condens. Matter **15**, L197 (2003).

[63] J. C. Thomas and A. Van der Ven, Phys. Rev. B **90**, 224105 (2014).

[64] J. G. Goiri and A. Van der Ven, Phys. Rev. B **94**, 094111 (2016).

[65] S. S. Behara and A. Van der Ven, Phys. Rev. Mater. **6**, 054403 (2022).

[66] S. K. Kolli, A. R. Natarajan, J. C. Thomas, T. M. Pollock, and A. Van der Ven, Phys. Rev. Mater. **4**, 113604 (2020).

[67] J. C. Thomas, A. R. Natarajan, and A. Van der Ven, npj Comput. Mater. **7**, 164 (2021).

[68] G. Ceder, Comput. Mater. Sci. **1**, 144 (1993).

[69] A. van de Walle and G. Ceder, Rev. Mod. Phys. **74**, 11 (2002).

[70] L. Chaput, A. Togo, I. Tanaka, and G. Hug, Phys. Rev. B **84**, 094302 (2011).

[71] K. Parlinski, Z. Q. Li, and Y. Kawazoe, Phys. Rev. Lett. **78**, 4063 (1997).

[72] D. Alfè, Comput. Phys. Commun. **180**, 2622 (2009).

[73] D. A. Kitchaev, I. J. Beyerlein, and A. Van der Ven, Phys. Rev. B **98**, 214414 (2018).

[74] J. W. Cahn and J. E. Hilliard, J. Chem. Phys. **28**, 258 (1958).

[75] J. W. Cahn, Acta Metall. **9**, 795 (1961).

[76] S. M. Allen and J. W. Cahn, Acta Metall. **27**, 1085 (1979).

[77] J. W. Cahn, Acta Metall. **10**, 907 (1962).

[78] J. C. Thomas and A. Van der Ven, Phys. Rev. B **88**, 214111 (2013).

[79] S. Rudraraju, A. Van der Ven, and K. Garikipati, Comput. Methods Appl. Mech. Eng. **278**, 705 (2014).

[80] V. Ozoliņš, C. Wolverton, and A. Zunger, Phys. Rev. B **57**, 6427 (1998).

[81] C. Wolverton, V. Ozoliņš, and A. Zunger, Phys. Rev. B **57**, 4332 (1998).

[82] C. Wolverton, V. Ozolins, and A. Zunger, J. Phys.: Condens. Matter **12**, 2749 (2000).

[83] J. Z. Liu and A. Zunger, Phys. Rev. B **77**, 205201 (2008).

[84] J. Z. Liu and A. Zunger, J. Phys.: Condens. Matter **21**, 295402 (2009).

[85] K. Wang, D. Cheng, and B.-C. Zhou, npj Comput. Mater. **9**, 75 (2023).

[86] C. Wolverton, Modell. Simul. Mater. Sci. Eng. **8**, 323 (2000).

[87] R. Drautz, Phys. Rev. B **99**, 014104 (2019).

[88] F. Musil, A. Grisafi, A. P. Bartók, C. Ortner, G. Csányi, and M. Ceriotti, Chem. Rev. **121**, 9759 (2021).

[89] V. L. Deringer, A. P. Bartók, N. Bernstein, D. M. Wilkins, M. Ceriotti, and G. Csányi, Chem. Rev. **121**, 10073 (2021).