Combined cluster and atomic displacement expansion for solid solutions and magnetism

Kevin F. Garrity[*]

Material Measurement Laboratory, National Institute of Standards and Technology, Gaithersburg, Maryland 20899, USA

Finite-temperature disordered solid solutions and magnetic materials are difficult to study directly using first-principles calculations, due to the large unit cells and many independent samples that are required. In this work, we develop a combined cluster expansion and atomic displacement expansion, which we fit to first-principles energies, forces, and stresses. We then use the expansion to calculate thermodynamic quantities at nearly first-principles levels of accuracy. Our model naturally includes both configurational and vibrational entropy, including anharmonic contributions and structural phase transitions. In addition, we can treat coupling between atomic displacement and chemical or magnetic degrees of freedom. As examples, we use our expansion to calculate properties of $Si_{1-x}Ge_x$, magnetic MnO, Al with vacancies, and $Ba_xSr_{1-x}TiO_3$. Finally, we demonstrate that by treating all the relevant degrees of freedom explicitly, we can in some cases achieve improved convergence of fitting parameters versus distance as compared to a pure cluster expansion.

I. INTRODUCTION

Solid solutions, which are materials that have well-defined crystal structures but disordered occupancy of atomic positions, are important for a variety of technological applications as both structural and functional materials [1]. About half of the Inorganic Crystal Structure Database consists of compounds with partial occupancy [2]. Similarly, compounds with spin degrees of freedom are often disordered at experimentally relevant temperatures. Unfortunately, both of these types of materials are difficult to treat at a first-principles level of accuracy. Large supercells and averages over many configurations are needed to treat disorder systematically, but the computational cost of plane-wave density functional theory (DFT) calculations increases rapidly with the number of atoms in a calculation [3]. Even worse, finite-temperature properties require averages over thousands of steps of atomic motion.

Cluster expansions, which consist of models where chemical or spin degrees of freedom are treated as interacting scalar variables on a lattice, with all atomic displacements relaxed, are widely used to map out the finite-temperature phase diagrams of alloys and solid solutions, as well as spin systems [4–8]. However, cluster expansions have several deficiencies that can limit their applicability. Most importantly, because they do not explicitly include structural relaxations, cluster expansions can only calculate a limited number of properties. In particular, interactions between atomic displacement and chemical degrees of freedom cannot be treated easily, which excludes technologically relevant materials properties like piezoelectrics, ferroelectrics, ferroelastics, magnetocalorics, etc. that involve coupled degrees of freedom. Similarly, because they eliminate all atomic degrees of freedom, cluster expansions do not capture the effects of vibrational free

energy, and attempts to add vibrational free energy can be computationally expensive [5,9,10]. Finally, the atomic displacement relaxations needed to fit cluster expansions are computationally expensive, as they require calculating the energies, forces, and stresses of many intermediate structures during a relaxation, but only the final energy is used in the model. This implicit treatment of relaxation can lead to effectively long-range and high-order interactions between cluster variables, even if the underlying physical interactions are short-ranged, making the fitting process inefficient [11–14].

In this work, we combine the framework of a cluster expansion for chemical or spin degrees of freedom with an atomic displacement expansion, which has long been used to calculate finite-temperature properties of crystalline materials. Atomic displacement expansions up to harmonic order, i.e., phonon calculations, are routinely done either as finite differences calculations or using DFT perturbation theory [15], and higher-order calculations are used to treat anharmonic properties like thermal conduction or phase transitions [16–22]. In contrast to some similar works on model Hamiltonians [23–25] that are sometimes used to treat solid solutions [26,27], in this work, we keep all atomic displacement degrees of freedom rather than only those related to soft modes. By including all atomic displacement degrees of freedom rather than only soft modes, our model can quantitatively calculate materials properties like thermal expansion, thermal conductivity, dielectric response, etc. that receive contributions from modes beyond a limited set of pre-selected soft modes.

By combining a cluster expansion with an atomic displacement expansion, including interactions between them, we get a model with many desirable properties. First, the combined expansion inherits many of the useful properties of separate atomic displacement and cluster expansions; namely, the model can be fit using standard linear least-squares fitting techniques, is systematically improvable, and can treat chemical disorder, vacancies, and magnetic disorder. Second, unlike models with specific physically inspired energy terms

*kevin.garrity@nist.gov

[28–30], our expansion can be easily applied to any crystal structure and applies equally well to metals, semiconductors, and ionic materials. Third, the model can be coupled to external fields in a natural manner. Finally, because the model uses all of the energies, forces, and stresses from any reference calculation and makes the relevant degrees of freedom explicit, it can in some cases be fit with even fewer DFT calculations than a pure cluster expansion that only considers chemical degrees of freedom.

We have made the code to fit our expansion to first-principles calculations and evaluate new structures available online at [31]. The rest of the work is organized as follows. In Sec. II, we describe the expansion form, symmetry properties, and fitting procedure we use in this work. In Sec. III, we fit the model to several example systems: $Si_{1-x}Ge_x$, magnetic MnO, Al with vacancies, and $Ba_xSr_{1-x}TiO_3$. Finally, in Sec. IV, we present our conclusions.

II. EXPANSION FORM AND FITTING

A. Expansion

Our model consists of a Taylor expansion around a high-symmetry reference structure in terms of both scalar degrees of freedom and vector atomic displacements, including interactions terms between them. In this section, we will treat the case of a solid solution with a single type of atomic substitution, which is represented by a scalar degree of freedom. Magnetic and vacancy cases will be examined in the following sections. In our formulas, the subscript indices i, j, ... run over the atomic sites of the high-symmetry supercell that we expand around, and the superscript indices x, y,... run over Cartesian directions.

Our model consists of three main terms: a cluster expansion, an atomic displacement expansion, and interaction terms between the two:

$$E_{\text{tot}} = E_{\text{cluster}} + E_{\text{atom}} + E_{\text{inter}}. \tag{1}$$

The form of the cluster expansion is well-known:

$$E_{\text{cluster}} = \sum_i J_i s_i + \frac{1}{2!} \sum_{ij} J_{ij} s_i s_j + \frac{1}{3!} \sum_{ijk} J_{ijk} s_i s_j s_k + \dots, \tag{2}$$

where $s_i = 0, 1$ are scalar degrees of freedom at site i, with 1 corresponding to a dopant atom being present, and J_{ij}, etc., represent fitting coefficients. In contrast to a normal cluster expansion, these energy terms represent the energy of dopant atoms in the unrelaxed high-symmetry reference structure, not relaxed structures. Instead, we treat atomic displacements explicitly as follows:

$$E_{\text{atom}} = \frac{1}{2!} \sum_{ij}^{xy} K_{ij}^{xy} u_i^x u_j^y + \frac{1}{3!} \sum_{ijk}^{xyz} K_{ijk}^{xyz} u_i^x u_j^y u_k^b + \dots, \tag{3}$$

where u_i^x is the displacement of atom i in direction x from its reference position. K_{ij}^{xy} is the fitting coefficient for the interaction between atom i moving in the x direction and atom j moving in the y direction; other terms are similar. There is no first-order term because we assume the high-symmetry structure is in equilibrium. The second-order term is the

standard harmonic force constant matrix, and higher-order terms are anharmonic force constants. Forces are obtained by taking a derivative with respect to u_i^x in the normal fashion: $F_i^x = -\partial E_{\text{tot}}/\partial u_i^x$.

Finally, we include interaction terms between the scalar and vector degrees of freedom:

$$E_{\text{inter}} = \sum_{ij}^x M_{ij}^x s_i u_j^x + \frac{1}{2!} \frac{1}{1!} \sum_{ijk}^x M_{ijk}^{\cdot x} s_i s_j u_k^x$$

$$+ \frac{1}{1!} \frac{1}{2!} \sum_{ijk}^{xy} M_{ijk}^{xy} s_i u_j^x u_k^y + \dots, \tag{4}$$

where M_{ij}^x, etc. are fitting coefficients for the interaction terms. For example, the first-order term in this expansion, with coefficient M_{ij}^x, is turned on if there is a dopant at site i ($s_i = 1$), and determines the forces on the surrounding atoms j in direction x that result from that substitution. Similarly, the term M_{ijk}^{xy} represents the change the spring constant between the atoms at sites j, k in directions x, y due to a dopant at site i.

This expansion is very general and can in principle be used for any combination of substitutions and atomic distortions that maintains the topology of the bonding in the crystal structure. While the expansion must be truncated in practice, it can be systematically improved if higher precision is needed. We will demonstrate in Sec. III that is also useful in practice, and as discussed in Sec. III A, this expansion will often have better convergence properties than an expansion that treats some of the degrees of freedom implicitly.

B. Symmetry

While the above expansion can in principle handle any reasonably small distortion of a unit cell, the number of fitting coefficients increases rapidly as higher-order terms are needed. To make the scheme useful, it is necessary to take advantage of symmetries of the reference structure in order to reduce the number of independent fitting coefficients. We will present a brief overview of the symmetry properties; most properties carry over from discussions of atomic displacement expansions [16,32].

The energy must be invariant under the application of the space group symmetries of the high-symmetry reference structure, which consist of a symmetry matrix R^{xy} and potentially a partial translation τ^x. Under the application of a symmetry operation, the site i can be shifted to another site i': $X_{i'}^x = \sum_y R^{xy} X_i^y + \tau^x$, where X_i^x is the reference position of atom i. Because the invariance must hold for any combination of s_i and u_i^x, each term in our expansion must be individually invariant. For example,

$$J_{ij} = J_{i'j'}, \tag{5}$$

$$K_{ij}^{xy} = \sum_{zw} R^{xz} R^{yw} K_{i'j'}^{zw}, \tag{6}$$

$$M_{ijk}^{\cdot x} = \sum_y R^{xy} M_{i'j'k'}^{\cdot y}. \tag{7}$$

These relations are widely known and used for separate cluster and atomic displacement expansions, and aside from keeping track of which degrees of freedom transform as

scalars and which as vectors, the relations in this work are analogous.

In addition to space group operations, the energy must be invariant under permutations of either the cluster degrees of freedom or the displacement degrees of freedom[16,32]. For example,

$$J_{ij} = J_{ji}, \tag{8}$$

$$K_{ij}^{xy} = K_{ji}^{yx}, \tag{9}$$

$$M_{ijk}^{\ x} = M_{jik}^{\ x}. \tag{10}$$

We note that scalar and vector degrees of freedom cannot be permuted for each other.

In addition to space group operations, each term in the model must also be invariant under arbitrary translations of the unit cell [16,32]. These relations are also known as acoustic sum rules because of their role in ensuring that there are three zero frequency phonon modes at Γ. The acoustic sum rules for our expansion are again simple generalizations of the relations from pure atomic displacement expansions. For example,

$$0 = \sum_j K_{ij}^{xy} \quad \forall i, xy, \tag{11}$$

$$0 = \sum_j M_{ij}^{\ x} \quad \forall i, x, \tag{12}$$

$$0 = \sum_k M_{ijk}^{xy} \quad \forall ij, xy. \tag{13}$$

These constraints are enforced as linear constraints during the fitting procedure.

In addition to the constraints discussed above, we note that there are similar constraints due to the invariance of the system under arbitrary rotations of the unit cell [16,32]. These additional constraints relate different orders of the expansion to each other; however, we do not enforce them explicitly during our fitting procedure.

C. Strain

In addition to cluster and atomic displacement variables, it is necessary to include strain degrees of freedom, ϵ^{xy}, in our model. However, strains are fundamentally related to long wavelength atomic displacements, and our existing expansion does not require any new fitting coefficients to treat strain [32,33]. The relationship between the harmonic force constants and the elastic constants $C^{wx,yz}$ is well-known, albeit rarely used in first-principles contexts:

$$E_{\text{strain}} = \frac{1}{2} \sum^{wx,yz} C^{wx,yz} \epsilon^{wx} \epsilon^{yz}, \tag{14}$$

$$S^{wx,yz} = \frac{1}{2} \sum_{ij} K_{ij}^{wx} (X_i^y - X_j^y)(X_j^z - X_i^z), \tag{15}$$

$$C^{wx,yz} = S^{wy,xz} + S^{xy,wz} - S^{xw,zy}, \tag{16}$$

where X_i^x is the reference position of atom i in direction x, and $S^{wx,yz}$ is a tensor defined above. Elastic constants have an extra permutation relation, $C^{wx,yz} = C^{yz,wx}$, that in some cases results in an additional constraint on the force constants.

We enforce this relation by requiring that the spring constants obey the Kun-Huang condition, $S^{wx,yz} = S^{yz,wx}$ [32,33].

The contributions to the elastic constants due to dopants are treated using analogous formulas, except with s_i variables that turn on the extra contributions in the presence of dopants. For example,

$$E_{\text{cl-strain}} = \frac{1}{2} \sum_i^{wx,yz} C_i^{wx,yz} s_i \epsilon^{wx} \epsilon^{yz}, \tag{17}$$

$$S_i^{wx,yz} = \frac{1}{2} \sum_{jk} M_{ijk}^{wx} (X_j^y - X_k^y)(X_j^z - X_k^z), \tag{18}$$

$$C_i^{wx,yz} = S_i^{wy,xz} + S_i^{xy,wz} - S_i^{xw,zy}. \tag{19}$$

In addition to the above terms, which are second-order in strain, there are additional first order in strain effective interactions. The lowest-order interaction term between pure atomic displacements and strain, $E_{\text{at-strain}}$, is

$$E_{\text{at-strain}} = \sum_i^{xyz} T_i^{x,yz} u_i^x \epsilon^{yz} + \dots, \tag{20}$$

$$T_i^{x,yz} = \sum_j K_{ij}^{xy} (X_j^z - X_i^z), \tag{21}$$

where $T_i^{x,yz}$ is the first-order coupling between strain ϵ^{yz} and the atomic displacement u_i^x. We emphasize that $T_i^{x,yz}$ is fully determined by appropriate combinations of our existing coupling coefficients and is not an independent fitting parameter. Similarly, the lowest-order coupling between a cluster variable s_i and strain ϵ^{xy}, is a simple generalization of Eq. (21):

$$E_{\text{cl-strain}} = \sum_i^{xy} U_i^{xy} s_i \epsilon^{xy} + \sum_i^{xy} V_{ij}^{x,yz} s_i u_j^x \epsilon^{yz} + \dots, \tag{22}$$

$$U_i^{xy} = \sum_j M_{ij}^x (X_j^y - X_i^y), \tag{23}$$

$$V_{ij}^{x,yz} = \sum_j M_{ijk}^{xy} (X_k^z - X_j^z). \tag{24}$$

By including all terms of Eqs. (14)–(24), we include the effects of strain up to second order; we ignore higher-order strain contributions. Stress is calculated in the normal way, $\sigma_{xy} = -\frac{1}{V} \partial E_{\text{tot}} / \partial \epsilon_{ij}$.

D. Vacancies

We can represent vacancies (or interstitial atoms) using the same formalism as discussed above, with nonzero cluster degrees of freedom representing missing atoms instead of substituted atoms. The only difficulty is that unless we impose additional constraints, the energy will depend weakly on the displacement of a vacancy site, which is unphysical as there is no atom to displace. The necessary constraints simply require the force on any vacancy site to be exactly zero. These constraints enforce a cancellation between various terms in the model to ensure the energy does not depend on vacancy positions. These constraints can be constructed naturally by the same procedures used to setup the linear regression.

E. Magnetism

We can use our expansion to treat simple magnetic systems, with the spins represented by cluster variables. If the spins are limited to collinear up and and down spins, and have nearly constant magnitude, then our expansion still applies. The only difference is that the cluster variables become $s_i = \pm 1$, as in the Ising model. A collinear magnetic field can be treated as a chemical potential.

Unlike cluster variables, which represent different atoms, there is an additional symmetry in spin systems due to the invariance of the energy when flipping every spin ($s_i \rightarrow -s_i$). This symmetry requires that expansion terms have an even number of spin variables.

In many cases, an Ising-like model will be inadequate to treat the magnetic degrees of freedom. A natural expansion of this formalism is to use the more general Heisenberg model, where the spins are allowed to rotate in three dimensions and magnetic anisotropy can be included. A full model would require treating the spin degrees of freedom like vectors with constant magnitude, instead of scalars. However, for simple situations, it is possible to fit the expansion to simple collinear spins, but then allow the spins to rotate during calculations using the model. This can be done by treating interactions between pairs of scalar spins like a dot product ($s_i s_j \rightarrow \vec{s}_i \cdot \vec{s}_j$). Using this idea, we can fit our expansion to collinear spin data using the Ising-like expansion, but allow for Heisenberg-like spin-spin interactions when calculating finite-temperature properties.

F. Fitting

Despite its generality, our model is linear in the coupling coefficients, and can be fit using standard linear least-squares techniques. We first use Gaussian elimination to determine the minimum number of independent fitting parameters after applying Eqs. (5)–(10). We prepare a library of energies, forces, and stresses from a set of DFT calculations that are relevant to the desired application. Then, we fit the coefficients, enforcing the acoustic sum rules [Eqs. (11)–(13)] as linear constraints.

It is necessary to truncate the expansion in order to perform the fitting. We generally truncate terms that are less than third order in the atomic displacements and less than third order in the chemical degrees of freedom only due to limitations imposed by the supercell. In order to keep the number of terms limited to those mostly likely to be significant, we limit such terms to two-body interactions beyond third nearest neighbors. Similarly, we truncate higher-order terms after second or third nearest neighbors. If the model fails to reach the desired level of accuracy and if sufficient fitting data is available, these constraints can be relaxed and additional terms can be added to the fitting, at the cost of a more time consuming fitting step. Despite our efforts to limit the number of terms, this still leaves a large number of terms to fit. In order to search for a sparse solution, we use recursive feature elimination [34], with cross-validation to determine how many independent parameters to keep for optimum out-of-sample prediction. During each step of this algorithm, the smallest standardized coefficients are dropped and the model is refit. We find that this procedure is faster and more numerically stable than L_1 regularization, which has previously been used to search for sparse models in a similar context [18].

For insulating materials, electrostatic dipole-dipole interactions decay very slowly with distance, and it is necessary to handle these contributions separately from our fitting procedure. We use Born effective charges and the electronic dielectric constant, determined from DFT perturbation theory [15,35], to subtract the contribution of long-range electrostatic forces before fitting. We then add back the long-range contribution for predictions. For the current examples where atoms of the same valence are substituted, we assume the Born effective charges are not modified by substitutions, atomic displacements, or strain. The accuracy of this approximation will depend on the details of the system being studied. In the cases we consider, we find that it is sufficient; however, we are considering better approximations to the long-range electrostatic coupling for use in future work. We clarify that only the dipole-dipole interaction beyond our harmonic fitting range is affected, as inside the fitting range the short-range interactions will compensate for any errors in the electrostatic coupling.

In order to test and improve our models, we also implement the ability to use automatically run finite-temperature Monte Carlo sampling to generate new structures, perform new DFT calculations on these structures, and add the new data to the model in order to iteratively improve it. In particular, we run Monte Carlo at a temperature relevant to the system we are studying, but in a small enough unit cell to perform an additional DFT calculations. We also perform sampling on the chemical degrees of freedom if we expect those degrees of freedom to be in equilibrium. After sufficient sampling to reach equilibrium, we take the final structure and perform a new DFT calculation. This procedure both provides a source of relevant out-of-sample data to test our model, and new data to improve the model. We stop this procedure when our predictions match the new DFT calculations to sufficient accuracy for our application.

This iterative technique is especially useful when the initial model has a hidden unphysical instability versus atomic displacements for at least some chemical degrees of freedom [21]. Because the phase space of the model (atomic displacements, chemical arrangements, and strain) is high dimensional, it can be possible for the model to appear to be accurate for many testing configurations, but still have a hidden instability that causes the model to either move to an unphysical lower energy ground state structure or fail to have any bounded energy at all, resulting in an energy below the expected energy range. Finite-temperature sampling of such an unstable model will allow the model to discover this region of phase space with incorrect behavior and get stuck there. Our iterative fitting procedure will then automatically perform a new DFT calculation and add this new structure to our set of fitting data, which should repair the instability for the next iteration. We continue the iterating until no new instabilities develop and the model reaches the desired accuracy.

G. First-principles computational details

We perform first-principles DFT calculations [36,37] with a plane-wave basis set as implemented in QUANTUM ESPRESSO

[38] and using the GBRV high-throughput ultrasoft pseudopotential library [39,40]. We use a plane wave cutoff of 40 Ryd. We use the PBEsol exchange-correlation functional [41], which provides more accurate lattice constants than other generalized gradient approximation functionals. For Mn d states, we use DFT+U with $U = 3$ eV. We calculate Born effective charges and dielectric constants using DFT perturbation theory [15]. We take advantage of the computational efficiency of using nondiagonal supercells to calculate long-range two-body model terms [42].

H. Monte Carlo sampling

In order to determine finite-temperature properties using our model, we perform classical Monte Carlo sampling of the Boltzmann distribution using the METROPOLIS algorithm [43]. We use a local updating strategy, performing sweeps where we attempt to move each atom a random distance, the average magnitude of which is tuned during thermalization to achieve approximately 50% acceptance. Strain is treated in a similar fashion. In cases where we also allow the cluster variables to change, we perform grand canonical Monte Carlo [44], which involves including additional cluster variable sweeps using a single spin flip approach. In order to sample the grand canonical ensemble, we cycle through atomic displacement, strain, and cluster sweeps, accepting proposed moves with the Boltzmann probability. Because all of the energy terms in our model (except for long-range electrostatic interactions) are local, a single site updating strategy will generally be efficient, except possibly near a phase transition where degrees of freedom can become strongly correlated.

III. EXAMPLES

A. Si$_{1-x}$Ge$_x$

As our first example, we consider the technologically relevant solid solution Si$_{1-x}$Ge$_x$, in the diamond structure. The end members are the only thermodynamically stable phases at zero temperature.

We begin by investigating the effects of atomic relaxation on the convergence of interactions in Si$_{1-x}$Ge$_x$ [13]. We consider a $4 \times 4 \times 4$ cell of Si (128 atoms) with two Ge atoms substituted for two Si. In Fig. 1, we plot the energy difference between a given configuration of Ge atoms, as a function of distance between them, taking as the reference an isolated Ge atom, either relaxed or unrelaxed, in the same cell. In the case where all the atoms are fixed to their ideal positions, the first neighbor effective interaction between the Ge is large, but decays to zero rapidly as the distance increases. Even the second neighbor direct interaction is almost negligible. In contrast, when all the atoms in the cell are allowed to relax, the first neighbor interaction energy is smaller, but it barely decays with distance. A cluster expansion fit to the relaxed energies will have much worse convergence with distance than a fixed atom cluster expansion. By taking into account atomic displacements explicitly, we can take advantage of this improved convergence. We typically find that the limiting factor in our expansions is the convergence of the force constants, rather than the pure cluster terms. Also, we find that in the fixed atom case, three-body cluster interactions
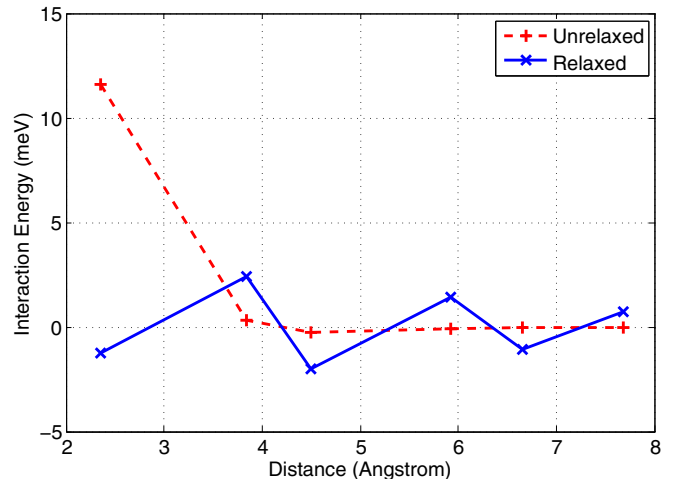

FIG. 1. Interaction energy between two Ge substituted for Si as a function of distance. Red dashed line is for unrelaxed positions, blue solid line is for relaxed positions.

are either very short-ranged or negligible. In Appendix, we discuss a direct comparison between a pure cluster expansion and a combined cluster and atomic displacement expansion, fit to the same DFT calculations.

Moving on to the fitting our expansion, we first fit to structures with only small atomic displacements. Our fitting data consists of structures with pure Si, pure Ge, and random substitutions of Si and Ge, as well as random atomic displacements of up to 0.15 Å. We use 40 calculations to fit, with supercell sizes of up to $4 \times 4 \times 4$ the primitive unit cell (128 atoms), and we test our expansion with another 30 structures, some of which have $2 \times 2 \times 8$ unit cells to test possible longer-range interactions. For this simple example case, we limit the range of the model to only second order in the cluster variables and up to third order in the atomic displacements, and the third-order terms are only nearest neighbor. We note that because the highest-order terms in our model are an odd power of the atomic displacements, the model cannot be applied for arbitrarily large displacements, but it can successfully reproduce the properties of the material within its range of validity.

After performing recursive variable selection and cross-validation, we are left with 95 independent fitting parameters out of 555 original parameters. We show the out-of-sample forces and energies predicted by the model in Fig. 2. We get excellent agreement over a considerable range of energies and forces, with a mean absolute error in energies of 0.2 meV/atom, or 1.0%, and a mean absolute error in force components of 1.4%.

Next, we generate structures with larger distortions, up to 0.7 Å, and we allow a more general set of anharmonic expansion terms, up to fourth order in the atomic displacements and out to second nearest neighbor, again using recursive feature elimination to identify the important coefficients. As shown in Fig. 3, our new model again has excellent performance, with an out-of-sample mean absolute error of 0.3 meV/atom, as compared to an average energy of 63 meV/atom in the testing data. The model is suitable for thermodynamic calculations
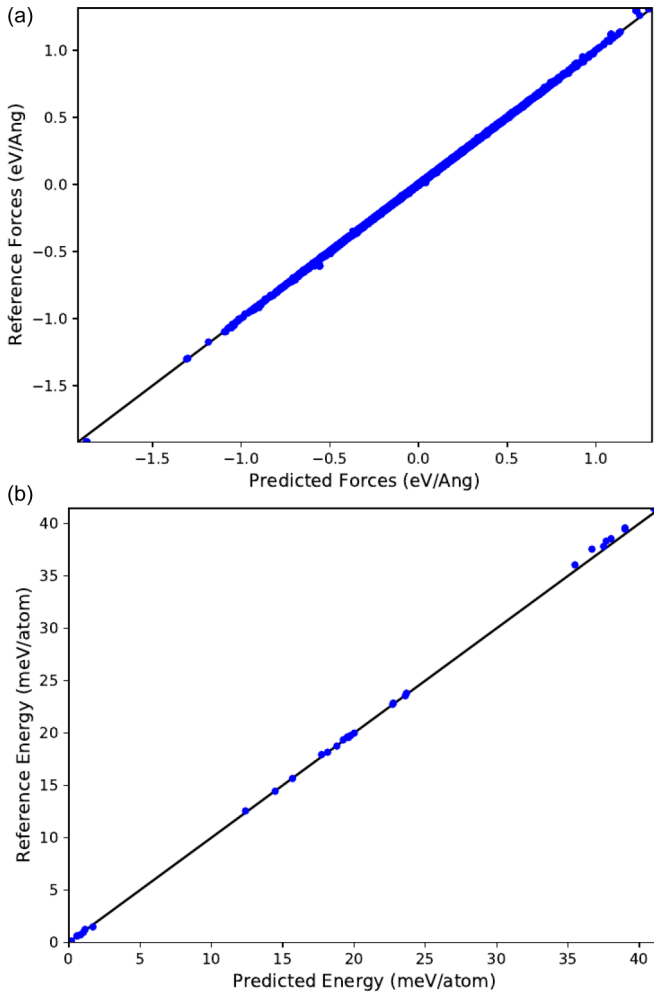
FIG. 2. For small distortions of the $Si_{1-x}Ge_x$ system, out-of-sample comparison between model (x axis) and first-principles (y axis) (a) forces (eV/Å) and (b) energies (meV/atom).

FIG. 3. For larger distortions of the $Si_{1-x}Ge_x$ system, out-of-sample comparison between model (x axis) and first-principles (y axis) (a) forces (eV/Å) and (b) energies (meV/atom).

up to several hundred Kelvin and includes all relevant anharmonic contributions to the energy.

In Fig. 4, we show an example grand canonical Monte Carlo calculation in a $10 \times 10 \times 10$ unit cell. Atomic displacements, strain, chemical degrees of freedom are all included in our grand canonical Monte Carlo sampling. For several fixed chemical potentials, we plot the Ge filling fraction as a function of temperature. At low temperature, the system is phase separated into pure Ge and pure Si. Near 300 K, depending on the chemical potential, the system transitions to a disordered solid solution, in reasonable agreement with past results on this system [6,13,45].

B. MnO

MnO in the rocksalt structure has an antiferromagnetic (AFM) ground state with a Neel temperature of 118 K [46]. Even without spin-orbit coupling, the spin structure breaks cubic symmetry and leads to a rhombohedral distortion of the unit cell of $0.96°$, according to our calculations. This distortion reduces the frustration of the first nearest neighbor antiferromagnetic spin-spin interactions, only half of which
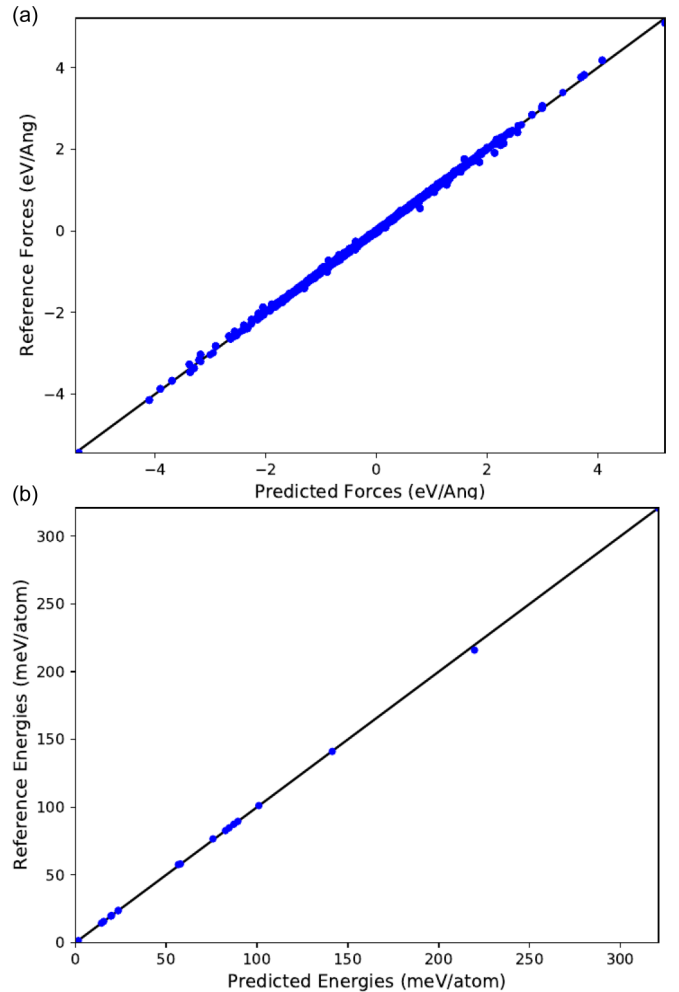
can be satisfied in the cubic structure [46]. The second neighbor interactions are comparable in size to the first neighbor interactions, and are not frustrated. We find that the energy difference between the ferromagnetic (FM) and antiferromagnetic phases in the cubic structure is 77 meV/Mn, and increases to 86 meV/Mn when the rhombohedral distortion is relaxed; so, the distortion energy is small, but not negligible.

We seek to study this coupled magnetic-structural phase transition using our model. We expand around the cubic ferromagnetic structure. We allow only second-order terms in the spin degrees of freedom, as discussed in Sec. II E. We consider first- and second-order atomic displacement terms, which can be coupled to spin, as well as short-range third and fourth-order displacement terms. We fit the model to a set of structures with both ordered and random collinear spins, and atomic displacements up to 0.5 Å in $4 \times 4 \times 4$ supercells. In Fig. 5, we show a comparison between our model and reference DFT calculations. We find excellent agreement, with a mean absolute energy error of 0.9 meV/Mn over a wide energy range.

We use the model to preform classical Monte Carlo sampling in a $12 \times 12 \times 12$ unit cell, interpreting the spins as
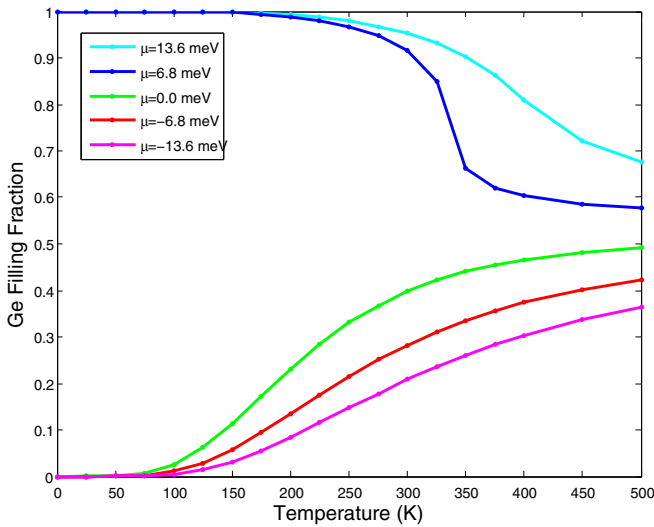
FIG. 4. Ge filling fraction as a function of temperature (K) for $Si_{1-x}Ge_x$, for fixed chemical potential μ. Different color lines are different chemical potentials, see legend.

Heisenberg-like, as discussed in Sec. II E. In Fig. 6(a), we plot the antiferromagnetic order parameter as a function of temperature, both using the full model and with the atoms fixed to the cubic structure and only the spins allowed to relax. The coupling between the spin variables and structural variables both raises the phase transition temperature and modifies its character, making the transition more strongly first-order. In addition, we can use the model to examine how the structure changes near the phase transition. In Fig. 6(a), we plot both the xx and xy components of strain as a function of temperature. As expected, the xy component is only nonzero in the low-temperature AFM phase, and its magnitude is closely related to the AFM order parameter. The xx component is less affected by the transition, but there is a minor change in slope as well as a small volume jump at the phase transition. However, this volume change is much smaller
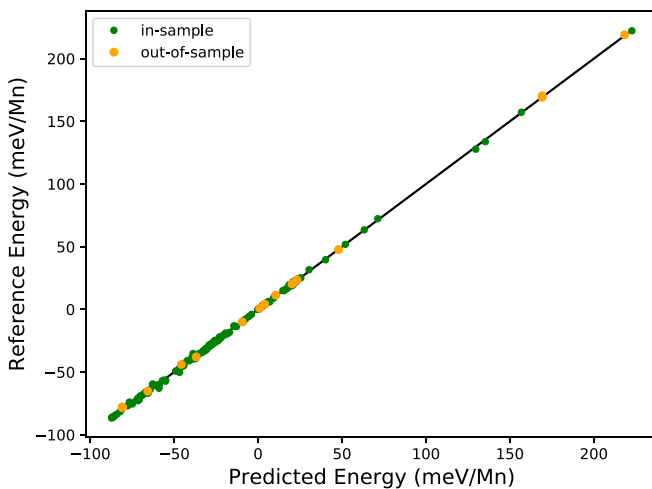

FIG. 5. For MnO, comparison between model energies and first-principles energies, in meV/Mn. Green symbols are in-sample, larger orange symbols are out-of-sample. Reference energy is cubic MnO in FM phase.
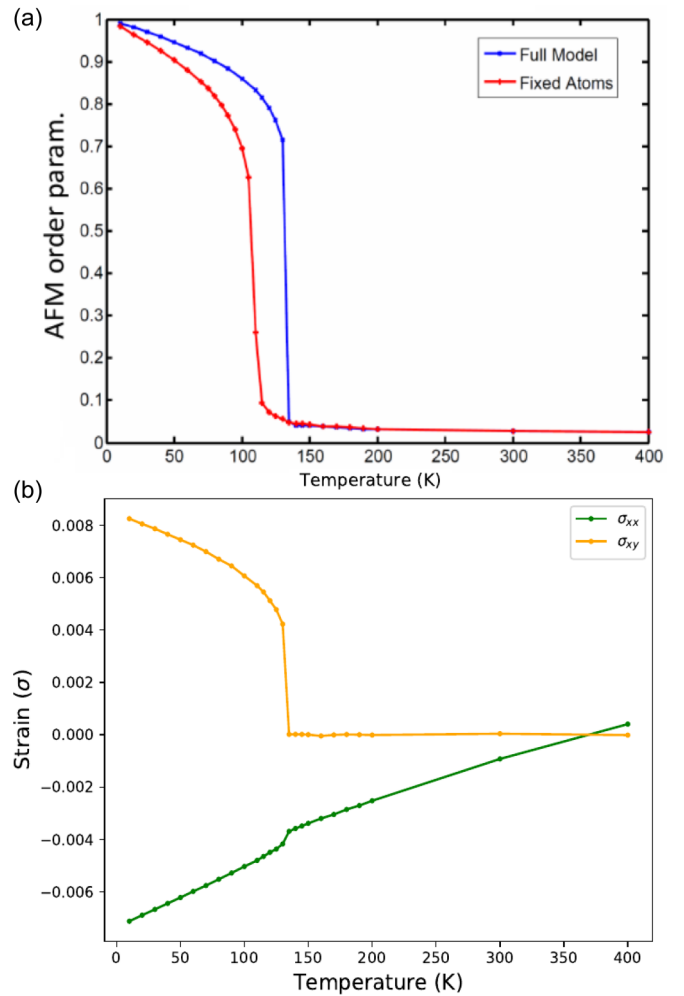

FIG. 6. (a) In MnO, AFM order parameter as a function of temperature (K). Blue line allows full atomic displacements, red line fixes them. (b) Strain vs temperature (K). Green line is σ_{xx}, orange is σ_{xy}.

than the 2.2% volume difference between the AFM and FM phases at zero temperature, which emphasizes the fact that the cubic paramagnetic phase at finite temperature is not well approximated by the cubic FM phase at zero temperature.

We note that the coupled structural-magnetic finite-temperature calculations performed in this section are straightforward using our expansion method, but would be challenging to calculate directly using first-principles techniques or using a pure spin-spin magnetic model.

C. Al with vacancies

As an example of a vacancy calculation, we consider Al in the *fcc* structure [47]. We consider $3 \times 3 \times 3$ and $6 \times 6 \times 6$ unit cells with 0–2 vacancies per cell (up to 7%), and atomic displacements of up to 0.26 Å. The reference chemical potential is that of zero-temperature bulk Al. Our expansion consists of terms up to fourth order in the atomic displacements and up two second order in the cluster variables, although the third- and forth-order terms are very short-ranged.

As shown in Fig. 7, our expansion gives excellent out-of-sample agreement with the reference forces and energies.
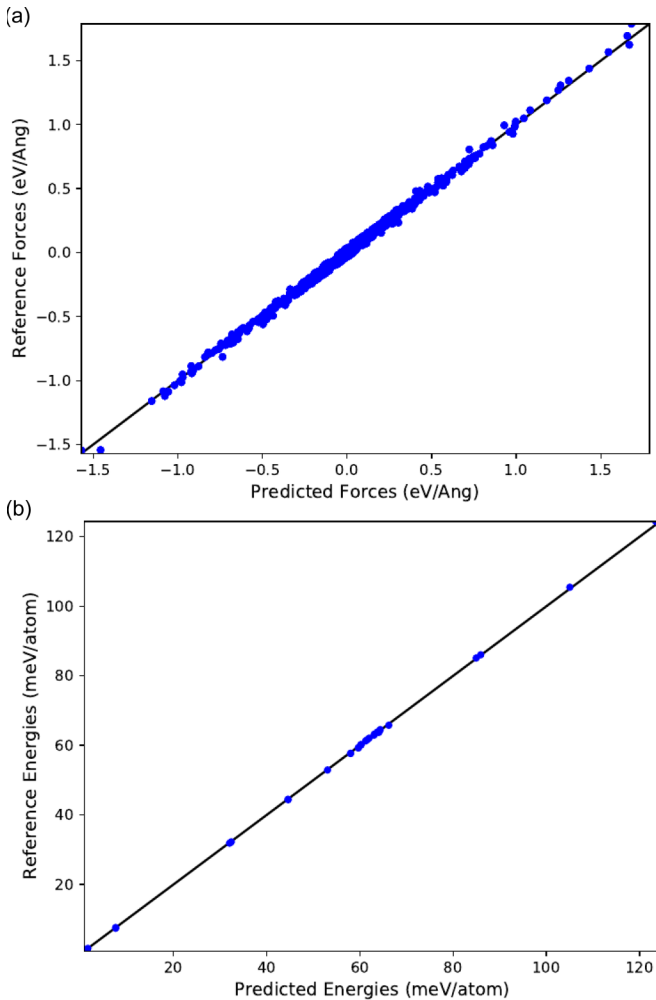
(a)

(b)

FIG. 7. For Al with vacancy system, out-of-sample comparison between model (x axis) and first- principles (y axis) (a) forces (eV/Å) and (b) energies (meV/atom).

One current limitation of the model is that it does not allow vacancies to hop from site to site, and it is not designed to reproduce this energy barrier between lattice sites, which is strongly anharmonic. This hopping will begin to happen at fairly low temperatures in Al. Because neither our expansion procedure nor our Monte Carlo routines are designed handle barrier hopping events, we are limited in the thermodynamic calculations we can perform on this system to low temperatures. Alternatively, we could limit our expansion to harmonic order in the atomic displacements, which would prevent hopping events at the cost of limiting our treatment of vibrational entropy to harmonic order. Any lattice-based expansion method will fail as the system approaches the melting temperature, but we expect that an extension of this calculational framework to handle rare barrier hopping events in a more systematic manner is possible, which we leave as a future direction of research.

D. $Ba_xSr_{1-x}TiO_3$

$BaTiO_3$ and $SrTiO_3$ are well-studied perovskite oxides that are used technologically for their dielectric properties

[19,21,23–25,27]. At low temperatures, $BaTiO_3$ is a rhombohedral ferroelectric, due to a polar distortion along the (111) direction. As the temperature is raised, $BaTiO_3$ becomes orthorhombic, with polarization along the (011) direction, then tetragonal, with polarization along the (001) direction, and finally cubic.

$SrTiO_3$ is also cubic at high temperatures, but goes through at a nonpolar tetragonal phase transition related to octahedral rotations ($a^0a^0c^-$ in Glazer notation [48]). As has been well-studied, at zero temperature in DFT calculations, $SrTiO_3$ still has a weak polar distortion, even after taking octahedral rotations into account [49]. Zero temperature quantum fluctuations, which we do not include in our model, are necessary to get the correct ground state of $SrTiO_3$, and are understood to be responsible for the enormous low temperature dielectric constant of $SrTiO_3$ [25].

We fit our expansion to the $Ba_xSr_{1-x}TiO_3$ system, expanding around cubic $BaTiO_3$. Due to the fact that both end members are unstable in their high-symmetry phases, this system requires a much more careful treatment of the anharmonic modes than previous examples. It is necessary to include DFT calculations from the various locally stable minima in addition to the experimentally observed structures in order to ensure that model gives good results at finite temperature. We find that we can get accurate results up to 300 K by expanding only up to fourth order in atomic displacements, second neighbor in distance, and including up to three-body interactions. We use the recursive approach discussed in Sec. II F to identify instabilities in the model and generate new fitting data until the model reaches sufficient accuracy. Including structures generated during iteration, a total of 250 DFT calculations were used to do the fitting, which is many more than was necessary for the previous examples. This total includes both relaxations to find the various local minima as well as single structure calculations to sample higher-temperature configurations. In addition, we weigh low-energy structures more in our fitting to ensure the local minima are well described.

As can be seen in Fig. 8, model performance is not quite as good as the previous examples that lack unstable modes, and it begins to degrade around 50 meV/atom. However, the mean absolute error in energy is still only 1.4 meV/atom, as compared to an average energy of 36.8 meV/atom in our test set. Furthermore, the model is more accurate at lower energies, allowing it to describe local minima correctly, which is necessary to capture low-temperature behavior.

We use our model to run Monte Carlo calculations at finite temperatures, with Sr and Ba distributed randomly, but fixed during each calculation, because we expect the chemical degrees of freedom to be quenched to a random configuration during the experimental crystal growth procedure, and therefore not in equilibrium at temperatures relevant to the ferroelectric transition. We use a $10 \times 10 \times 10$ unit cell. In Fig. 9, we show the resulting phase diagram. We reproduce the three ferroelectric phase transitions on the Ba-rich portion of the phase diagram, as well as the nonpolar phase transition on the Sr-rich side. In agreement with previous first-principles based calculations in this system [19,21,23–25,27], we overestimate the $SrTiO_3$ phase transition temperature and underestimate the $BaTiO_3$ phase transition temperatures. Considering the small

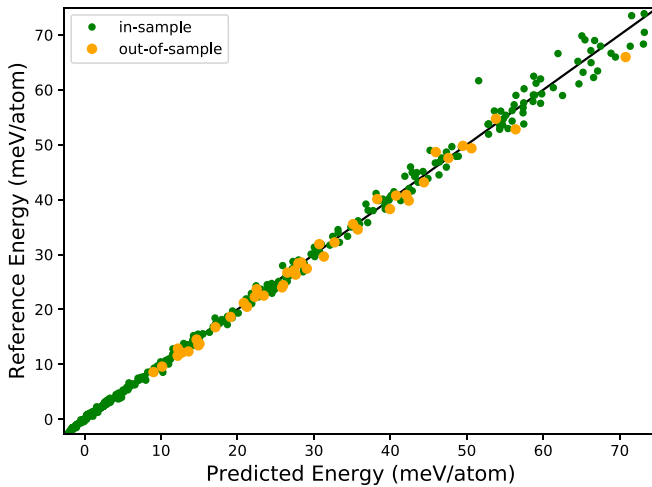FIG. 8. For $Ba_xSr_{1-x}TiO_3$, comparison between model energies and first-principles energies, in meV/atom. Green symbols are in-sample, larger orange symbols are out-of-sample.

FIG. 10. Static dielectric constant as a function of Ba percentage and temperature (K). Colors range from 0 (dark blue) to 2000 (dark red).

energy differences involved and the sensitivity of unstable modes to changes in volume, this level of accuracy is typical for first-principles phase diagrams. Due to the fact that we do not include quantum fluctuations, we find that the low temperature phase of $SrTiO_3$ is polar, instead of a quantum paraelectric.

In addition to identifying the phases, we can use our model to calculate detailed material properties as a function of doping and temperature. For example, in Fig. 10, we present the average dielectric constant throughout the phase diagram, calculated with the method of Ref. [26]. Unlike a effective Hamiltonian approach [26], our expansion includes contributions from all atomic degrees of freedom, instead of just soft modes, and treats thermal expansion correctly. Due limitations in converging the dielectric constant near a
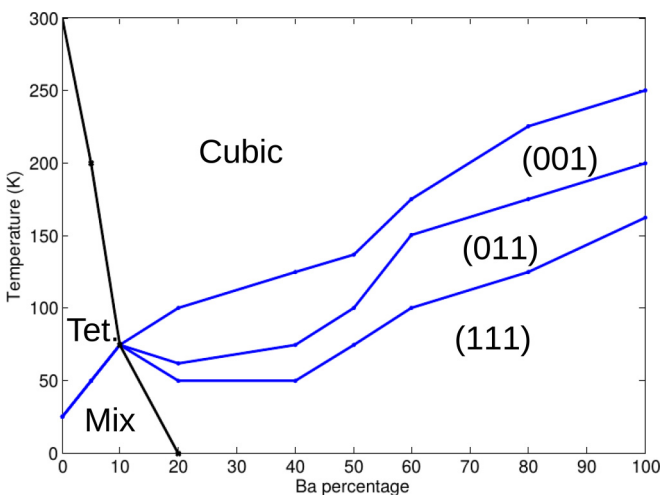
ferroelectric phase transition in a finite cell, we cap the reported dielectric constant at 2000. As expected, the dielectric constant is very high throughout the region where the various polar phase transitions occur, but is unaffected by the nonpolar transition in the Sr-rich region.

IV. CONCLUSIONS

In this work, we present an expansion in terms of both scalar degrees of freedom, corresponding to chemical or magnetic variables, and vector atomic displacements. We discuss various symmetry properties of the expansion, as well as a procedure for determining relevant coefficients and fitting them to first-principles calculations. By using examples, we show that the model can be usefully applied to solid solutions of semiconductors like $Si_{1-x}Ge_x$ and oxides like $Ba_xSr_{1-x}TiO_3$, as well as magnetic insulators like MnO and metals with vacancies like Al.

Due to the fact that this expansion can be applied to a wide range of materials, and that the fitting, improvement, and validation of model can proceed nearly automatically, we expect that it can be useful for many purposes. By combining features of a cluster expansion with structural degrees of freedom, we can in some cases achieve improved convergence with distance, and by making use of energies, forces, and stresses from every available self-consistent field calculation, we can fit the expansion with surprisingly few first-principles calculations. We fit our models using dozens of calculations for materials without structural transitions, and a few hundred for materials with multiple local minima and strong anharmonic terms. The expansion naturally includes both configurational free energy and vibrational free energy to any desired order. Most importantly, we can calculate properties that couple atomic displacements with chemical or magnetic degrees of freedom. This allows for the study of materials like ferroelectrics, piezoelectrics, electrocalorics, magnetocalorics, shape memory alloys, and ferroelastics that

FIG. 9. Phase diagram of $Ba_xSr_{1-x}TiO_3$ as a function of doping and temperature (K). Tet refers to nonpolar tetragonal, Mix refers to mixed octahedral rotations and polarization, and (111), (011), and (001) refer to polarization directions.
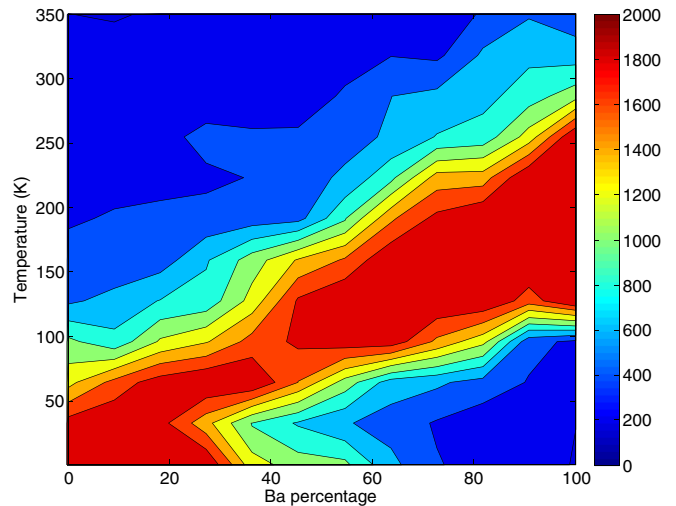
are technologically relevant but difficult to treat directly with first-principles calculations.

ACKNOWLEDGMENT

We wish to acknowledge discussions with E. Cockayne and K. Choudhary.

APPENDIX: CONVERGENCE VERSUS PURE CLUSTER EXPANSION

The primary purpose of the combined atomic displacement and cluster expansion presented in this work is not to compete directly with pure cluster expansion treatments of chemical disorder, but instead to calculate properties that depend on atomic displacements. Nevertheless, in some cases it is possible to take advantage of the shorter convergence length of the unrelaxed cluster expansion coefficients, as discussed in Sec. III A and Fig. 1, to fit a combined model that can out-perform a pure cluster expansion fit to the same number of DFT calculations. This out-performance with limited data is possible despite that fact the full model requires fitting many more total coefficients.

The counter-intuitive potential for a more complicated model to have improved performance despite limited fitting data depends on the fact that a pure cluster expansion can make use of only a very limited fraction of the total information produced during a typical first-principles structural relaxation, namely, the final energy. In contrast, the combined expansion presented in this work can take advantage of the energy, forces, and stress from every step of the structural relaxation. Therefore any given calculation with N atoms and n_s steps represents $n_s (3(N-1)+6+1)$ total pieces of fitting data. In practice, this data will be highly correlated during the final relaxation steps. Nevertheless, there is clearly more data available to the combined expansion, which can compensate for the larger number of coefficients that are needed. When this additional data is combined with the shorter convergence length for unrelaxed coefficients, improved performance relative to a pure cluster expansion is possible, although certainly not guaranteed.

We present one example of this improved convergence in Fig. 11. In this example for the Si_xGe_{1-x} system examined in Sec. III A, we fit each model to the same nine structural relaxations, with x varying from 0.008–1.0, and the structural relaxations starting with the unrelaxed diamond structure. We
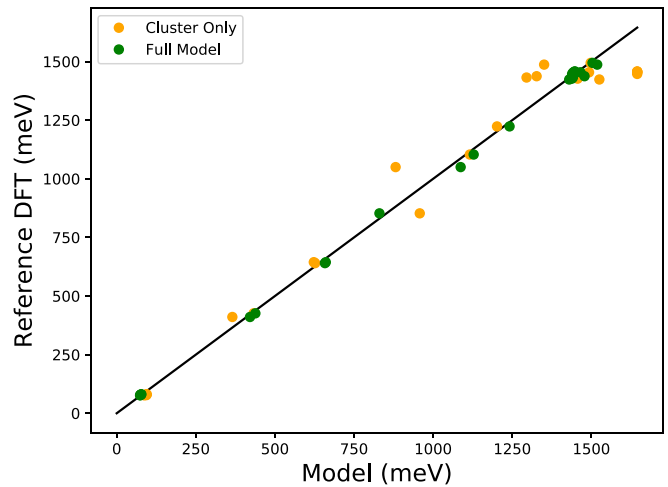

FIG. 11. Out-of-sample model comparison between model (x axis) and reference energy (y axis) of a pure cluster expansion model (orange points) and a combined cluster plus atomic displacement expansion (green points) for Si_xGe_{1-x} system as discussed in Appendix.

can see that although both models perform well given the limited fitting data, the combined model is more accurate than the pure cluster expansion in this case. Because the combined expansion also gives access to more properties than the cluster expansion, simply fitting the more complicated model with the same amount of data is already quite useful. We further clarify that this model fit on limited data is not expected to be as accurate as the model fit to more data in the main text.

In general, we do not expect either type of model to systematically outperform the other when fit to limited datasets, and we expect the same performance when each model has access to a sufficiently large dataset. Mostly likely, the pure cluster expansion will perform better when the effective cluster interactions converge quickly with distance, when the strain remains small, and when the atomic displacement expansion is long-ranged or high-order. In addition, we note that any Monte Carlo calculations will be less computationally expensive using the pure cluster expansion. Therefore a choice between the two models will most likely depend primarily on which properties are being studied, rather than the convergence. We leave a more systematic comparison between the two models to future work.

[1] W. D. Callister and D. G. Rethwisch, *Materials Science and Engineering: An Introduction*, 9th ed. (Wiley, 2013).

[2] A. Belsky, M. Hellenbrandt, V. L. Karen, and P. Luksch, Acta Crystallogr. Sec. B **58**, 364 (2002).

[3] A. Zunger, S.-H. Wei, L. G. Ferreira, and J. E. Bernard, Phys. Rev. Lett. **65**, 353 (1990).

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

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

[6] A. van de Walle and G. Ceder, J. Phase Equilib. **23**, 348 (2002).

[7] A. van de Walle, Calphad **33**, 266 (2009).

[8] Q. Wu, B. He, T. Song, J. Gao, and S. Shi, Comput. Mater. Sci. **125**, 243 (2016).

[9] G. D. Garbulsky and G. Ceder, Phys. Rev. B **49**, 6327 (1994).

[10] L. Anthony, J. K. Okamoto, and B. Fultz, Phys. Rev. Lett. **70**, 1128 (1993).

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

[12] V. Ozoliņš, C. Wolverton, and A. Zunger, Phys. Rev. B **57**, 4816 (1998).

[13] S. de Gironcoli, P. Giannozzi, and S. Baroni, Phys. Rev. Lett. **66**, 2116 (1991).

[14] M. Peressi and S. Baroni, Phys. Rev. B **49**, 7490 (1994).

[15] S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi, Rev. Mod. Phys. **73**, 515 (2001).

[16] K. Esfarjani and H. T. Stokes, Phys. Rev. B **77**, 144112 (2008).

[17] M. H. F. Sluiter, M. Weinert, and Y. Kawazoe, Phys. Rev. B **59**, 4100 (1999).

[18] F. Zhou, W. Nielson, Y. Xia, and V. Ozoliņš, Phys. Rev. Lett. **113**, 185501 (2014).

[19] J. C. Wojdeł, P. Hermet, M. P. Ljungberg, P. Ghosez, and J. Íñiguez, J. Phys.: Condens. Matter **25**, 305401 (2013).

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

[21] C. Escorihuela-Sayalero, J. C. Wojdeł, and J. Íñiguez, Phys. Rev. B **95**, 094115 (2017).

[22] W. Li, J. Carrete, N. A. Katcho, and N. Mingo, Comput. Phys. Commun. **185**, 1747 (2014).

[23] W. Zhong, D. Vanderbilt, and K. M. Rabe, Phys. Rev. Lett. **73**, 1861 (1994).

[24] W. Zhong, D. Vanderbilt, and K. M. Rabe, Phys. Rev. B **52**, 6301 (1995).

[25] W. Zhong and D. Vanderbilt, Phys. Rev. B **53**, 5047 (1996).

[26] L. Bellaiche, A. García, and D. Vanderbilt, Phys. Rev. Lett. **84**, 5427 (2000).

[27] L. Walizer, S. Lisenkov, and L. Bellaiche, Phys. Rev. B **73**, 144105 (2006).

[28] Y.-H. Shin, V. R. Cooper, I. Grinberg, and A. M. Rappe, Phys. Rev. B **71**, 054104 (2005).

[29] J. Tersoff, Phys. Rev. B **37**, 6991 (1988).

[30] M. Manninen, Phys. Rev. B **34**, 8486 (1986).

[31] https://github.com/usnistgov/spring_cluster

[32] G. Leibfried and W. Ludwig, in *Theory of Anharmonic Effects in Crystals*, Solid State Physics, edited by F. Seitz and D. Turnbull (Academic Press, New York, 1961), pp. 276–288.

[33] K. Huang, Proc. R. Soc. London A **203**, 178 (1950).

[34] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, Mach. Learn. **46**, 389 (2002).

[35] X. Gonze and C. Lee, Phys. Rev. B **55**, 10355 (1997).

[36] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[37] W. Kohn and L. Sham, Phys. Rev. **140**, A1133 (1965).

[38] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. D. Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, J. Phys.:Condens. Matter **21**, 395502 (2009).

[39] D. Vanderbilt, Phys. Rev. B **41**, 7892 (1990).

[40] K. F. Garrity, J. W. Bennett, K. M. Rabe, and D. Vanderbilt, Comput. Mater. Sci **81**, 446 (2014).

[41] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke, Phys. Rev. Lett. **100**, 136406 (2008).

[42] J. H. Lloyd-Williams and B. Monserrat, Phys. Rev. B **92**, 184301 (2015).

[43] W. K. Hastings, Biometrika **57**, 97 (1970).

[44] D. Frenkel and B. Smit, in *Understanding Molecular Simulation*, 2nd ed. edited by D. Frenkel and B. Smit (Academic Press, San Diego, 2002), pp. 1–6.

[45] A. Qteish and R. Resta, Phys. Rev. B **37**, 6983 (1988).

[46] J. A. M. Paddison, M. J. Gutmann, J. R. Stewart, M. G. Tucker, M. T. Dove, D. A. Keen, and A. L. Goodwin, Phys. Rev. B **97**, 014429 (2018).

[47] D. E. Turner, Z. Z. Zhu, C. T. Chan, and K. M. Ho, Phys. Rev. B **55**, 13842 (1997).

[48] A. M. Glazer, Acta Crystallogr., Sec. B **28**, 3384 (1972).

[49] N. Sai and D. Vanderbilt, Phys. Rev. B **62**, 13942 (2000).