Anharmonic free energy of lattice vibrations in fcc crystals from a mean-field bond

Thomas D. Swinburne [1,*] Jan Janssen [2] Mira Todorova [2] Gideon Simpson [3] Petr Plechac,[4]
Mitchell Luskin,[5] and Jörg Neugebauer [2]

[1]*Aix-Marseille Université, CNRS, CINaM UMR 7325, Campus de Luminy, 13288 Marseille, France*
[2]*Department of Computational Materials Design, Max-Planck-Insitut für Eisenforschung GmbH,
Max-Planck-Strasse 1, D-40237 Düsseldorf, Germany*
[3]*Department of Mathematics, Drexel University, Philadelphia, Pennsylvania 19104, USA*
[4]*Department of Mathematical Sciences, University of Delaware, Newark, Delaware 19716, USA*
[5]*School of Mathematics, University of Minnesota, Minneapolis, Minnesota 55455, USA*

It has recently been shown that the *ab initio* anharmonic free energy of fcc crystals can be approximated to meV/atom accuracy by a lattice of anharmonic nearest-neighbor bonds, where the bonding potential can be efficiently parametrized from the target system. We develop a mean-field approach for the free energy of a general bond lattice, analytically accounting for strong bond-bond correlations while enforcing material compatibility and thermodynamic self-consistency. Applying our fundamentally anharmonic model to fcc crystals yields free energies within meV/atom of brute force thermodynamic integration for core seconds of computational effort. Potential applications of this approach in computational materials science are discussed.

Introduction. The key quantity to determine material properties at finite temperature is the free energy [1,2]. Compared to $T = 0$ K calculations, where only a single minimized configuration is required, free-energy computation requires thermodynamic ensemble averages consisting of millions of individual configurations [3,4]. Computing such large numbers of configurations is impractical with *ab initio* calculations, but these are required to guarantee accuracy. As a result, to compute free energies the majority of current studies have to employ severe approximations. The most widely used are the harmonic approximation and its variant, the quasi-harmonic approximation (QHA), where a single anharmonic degree, the volume expansion, is included [3,5,6]. The reason for this popularity is the ease of construction and the existence of analytical expressions which directly connect the eigenvalues of the dynamical matrix with the free energy [7,8].

Recent studies [9,10] show, however, that the QHA may fail and provide even qualitatively wrong thermodynamic trends, motivating numerous approaches over the past few years to explicitly account for anharmonicity beyond the QHA [10–14]. An almost universal prerequisite for these methods is knowledge of a reference system where the free energy can be easily computed. Examples of such methods are thermodynamic integration or free-energy perturbation [4]. Due to the existence of an analytical expression the QHA is typically the preferred choice, but a critical requirement for such methods to be numerically efficient is that the potential energy surface of the reference system closely matches that of the true system in the region of configuration space where thermodynamic sampling occurs. Wide-range studies, e.g., on fcc metals [10]

showed that for many of the studied systems the conventional QHA is insufficient to allow accurate *ab initio* free-energy computation. When analyzing the reason Glensk *et al.* [10] discovered that a simple anharmonic nearest-neighbor bond model with only one (longitudinal) or two (+ transversal) force components provides a reference system that matches *ab initio* data with meV/atom precision. Thus, a lattice of anharmonic bonds allows one to describe the thermodynamic configuration space with very few and easy to determine parameters.

To fully exploit these insights we require methods that allow us to compute free energies of anharmonic bond models with the same computational efficiency and conceptual ease as of harmonic ones, allowing access to anharmonic free energies with negligible computational cost.

In this Rapid Communication we introduce an analytic model that fulfills this aim, giving anharmonic free energies for a model fcc bond lattice within meV/atom to brute force sampling, even for highly anharmonic systems. The computational effort is negligible for a given bonding potential, as no sampling is required. The present approach can thus be used in a variety of settings in multiscale materials modeling, some of which we discuss in the Conclusions. The PYTHON/C++ code used to generate reference data and evaluate model predictions is freely available [15], while data analysis and model predictions can also be run in the cloud [16].

Thermodynamics of a bond lattice. Consider a crystalline solid of N atoms in d dimensions, with ionic positions $\{\mathbf{R}_i\}_{i=1}^{i=N}$, confined to a fixed volume $(1 + \epsilon)V_0$ such that $\epsilon = 0$ corresponds to zero strain at zero temperature. The bond lattice model builds a surrogate potential energy using $mN/2$ pairwise interactions between an atom i and its m neighbors $n_l(i)$, where $l \in [1, m]$. This naturally motivates the definition
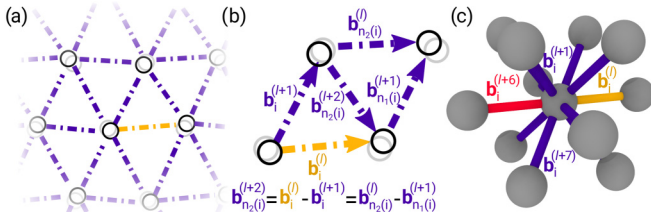
FIG. 1. (a) Cartoon of a hexagonal bond lattice with $m = 6$ bond vectors per atom. (b) The corresponding connectivity condition on the bond vectors, which in general is given by (3). (c) Labeling convention for the fcc lattice with $m = 12$ bond vectors per atom.

of mN "bond vectors"

$$\mathbf{b}_i^{(l)} \equiv \mathbf{R}_{n_l(i)} - \mathbf{R}_i \in \mathbb{R}^d. \qquad (1)$$

With a potential energy function $V_l(\mathbf{b}^{(l)})$ for a bond l the total energy reads

$$U(\{\mathbf{b}\}) = \frac{1}{2} \sum_{i=1}^{N} \sum_{l=1}^{m} V_l(\mathbf{b}_i^{(l)}), \qquad (2)$$

where the factor of $1/2$ corrects for the fact that two neighboring atoms share the same bond. The set of all bond vectors is clearly a heavily overcomplete description of the system, as each d-dimensional atom coordinate $\mathbf{R}_i$ has been replaced with m d-dimensional bond vectors $\{\mathbf{b}_i^{(l)}\}_{l=1}^{l=m}$. The overcompleteness of bond space is lifted by enforcing material compatibility, defined as a dN-dimensional surface $\mathcal{C}(\{\mathbf{b}\}) = 0$ in the mdN-dimensional bond space, corresponding to bond configurations satisfying (1). Material compatibility requires that elementary closed circuits around adjacent bonds give zero net displacement [17]. To define $\mathcal{C}(\{\mathbf{b}\}) = 0$, we first use the pairwise symmetry of (1) to set $\mathbf{b}_i^{(l)} = -\mathbf{b}_{n_l(i)}^{(l+m/2)}$ with no loss of generality. It is then simple to show that compatibility imposes the constraints

$$\mathbf{b}_i^{(l)} - \mathbf{b}_{n_k(i)}^{(l)} = \mathbf{b}_i^{(k)} - \mathbf{b}_{n_l(i)}^{(k)}, \qquad (3)$$

which defines the dN-dimensional constraint surface $\mathcal{C}(\{\mathbf{b}\}) = 0$ [18]. In Fig. 1(b) we provide a demonstration of the relations that lead to (3) for the two-dimensional hexagonal lattice. The equilibrium distribution in the mdN-dimensional bond space is then formally given by

$$\rho(\{\mathbf{b}\}, \epsilon, T) = \exp\left[-U(\{\mathbf{b}\})/(k_B T)\right]\delta[\mathcal{C}(\{\mathbf{b}\})]/Z. \qquad (4)$$

As the total energy (2) is a sum over bonds, the anharmonic contribution to the internal energy per atom at a temperature T and strain ϵ reduces to

$$\mathcal{U}^{\mathrm{ah}}(T, \epsilon) \equiv \frac{1}{2} \sum_l \int V_l(\mathbf{b}) \rho_{1,l}(\mathbf{b}, \epsilon, T) d\mathbf{b} - \frac{3}{2} k_B T, \qquad (5)$$

where we have defined the *single bond density* for a bond l in d-dimensional space at a temperature T and strain ϵ,

$$\rho_{1,l}(\mathbf{b}, \epsilon, T) \equiv \int \delta(\mathbf{b}_j^{(l)} - \mathbf{b}) \rho(\{\mathbf{b}\}, \epsilon, T) \prod_{il} d\mathbf{b}_i^{(l)}, \qquad (6)$$

which is identical for any choice of $j \in [1, N]$. Using elementary thermodynamic identities, the anharmonic component of

the free energy is then obtained from the integral

$$\mathcal{F}^{\mathrm{ah}}(T, \epsilon) \equiv T \int_0^T \mathcal{U}^{\mathrm{ah}}(T', \epsilon)/(T')^2 dT'. \qquad (7)$$

As one can access the internal energy and thus the free energy of a bond lattice through single bond averages, i.e., without any sampling, approximation of the singlebond densities $\rho_{1,l}(\mathbf{b}, \epsilon, T)$ is the central topic of this contribution.

Bond lattice for a model fcc crystal. While the above theoretical considerations are general to any crystal structure, in the present work we focus on fcc lattices, a close-packed structure which is common to many material systems and can be stabilized using a single bonding potential. The bond network is constructed from $m = 12$ nearest-neighbor bonds, with a labeling convention shown in Fig. 1(c). As all bonds are equivalent (reducible) under the fcc space group, in the absence of lattice vibrations a bond vector $\mathbf{b}^{(l)}$ is given by $\mathbf{a}_l^\epsilon \equiv (1 + \epsilon)\mathbf{a}_l$, where $\mathbf{a}_l$ is a primitive lattice vector at $\epsilon = 0$. Any two primitive lattice vectors $\mathbf{a}_l^\epsilon$, $\mathbf{a}_k^\epsilon$ satisfy $\mathbf{a}_l^\epsilon = \mathbf{G}_{lk}\mathbf{a}_k^\epsilon$, where $\mathbf{G}_{lk}$ represents a point symmetry. For the fcc lattice we can thus write $V_l(\mathbf{b}^{(l)}) = V_1(\mathbf{G}_{1l}\mathbf{b}^{(l)})$ and $\rho_{1,l}(\mathbf{b}^{(l)}, \epsilon, T) = \rho_1(\mathbf{G}_{1l}\mathbf{b}^{(l)}, \epsilon, T)$ giving a total energy

$$U_{\mathrm{fcc}}(\{\mathbf{b}\}) = \frac{1}{2} \sum_{i=1}^{N} \sum_{l=1}^{12} V_1(\mathbf{G}_{1l}\mathbf{b}_i^{(l)}), \qquad (8)$$

and an anharmonic internal energy per atom

$$\mathcal{U}_{\mathrm{fcc}}^{\mathrm{ah}}(T, \epsilon) = 6 \int \rho_1(\mathbf{b}, \epsilon, T) V_1(\mathbf{b}) d\mathbf{b} - \frac{3 k_B T}{2}. \qquad (9)$$

To generate reference data we used the "polarized" bonding potential derived from *ab initio* molecular dynamics (MD) for fcc bulk Al [10],

$$V_1(\mathbf{b}) \equiv V_\parallel(\hat{\mathbf{a}}_1 \cdot \mathbf{b}) + V_\perp(|\hat{\mathbf{a}}_1 \times \mathbf{b}|), \qquad (10)$$

where $\hat{\mathbf{a}}_1 = \mathbf{a}_1^\epsilon/|\mathbf{a}_1^\epsilon|$ and $|\hat{\mathbf{a}}_1 \times \mathbf{b}|$ projects the component of $\mathbf{b}$ perpendicular to $\hat{\mathbf{a}}_1$. The longitudinal potential $V_\parallel(b_\parallel) = D(1 - \exp[-\alpha(b_\parallel - a)])^2$ is of the Morse form, and captures the strong anharmonicity in the interatomic interactions. The transverse potential $V_\perp(b_\perp) \equiv \kappa D\alpha^2 b_\perp^2$ is harmonic, where κ controls the relative strength of the transverse potential. The potential parameters were $D = 0.1$ eV, $\alpha = 1.5$ Å^{-1}, $a = 2.85$ Å, with a variable transverse strength $\kappa \in [0.0, 0.3]$. A plot of $V_1(\mathbf{b})$ for $\kappa = 0.2$ is given in Fig. 2, where the significant anharmonicity is clear.

Sampling bond space under $\mathcal{C}(\{\mathbf{b}\}) = 0$ is equivalent to sampling position space while extracting bond configurations through (1). Reference data to benchmark against our sampling-free bond model was thus collated from independent MD trajectories governed by Eq. (8). Integration was performed with an overdamped Langevin thermostat. Thermalization of the system was carefully monitored by ensuring the measured virial temperature $T_{\mathrm{vir}}^{\mathrm{MD}} \equiv \langle \mathbf{R} \cdot \nabla U \rangle/(3k_B)$ agreed with the target temperature, i.e., $T_{\mathrm{vir}}^{\mathrm{MD}} = T$. Following thermalization, at each timestep the bond energy $V_1(\mathbf{G}_{1l}\mathbf{b}_i^{(l)})$ and bond vector $\mathbf{G}_{1l}\mathbf{b}_i^{(l)}$ was collated from the MD runs to estimate $\mathcal{U}^{\mathrm{ah}}(T, \epsilon)$ and $\rho_1(\mathbf{b}, \epsilon, T)$. This results in excess of 10^9 samples of the bond configurations, requiring around
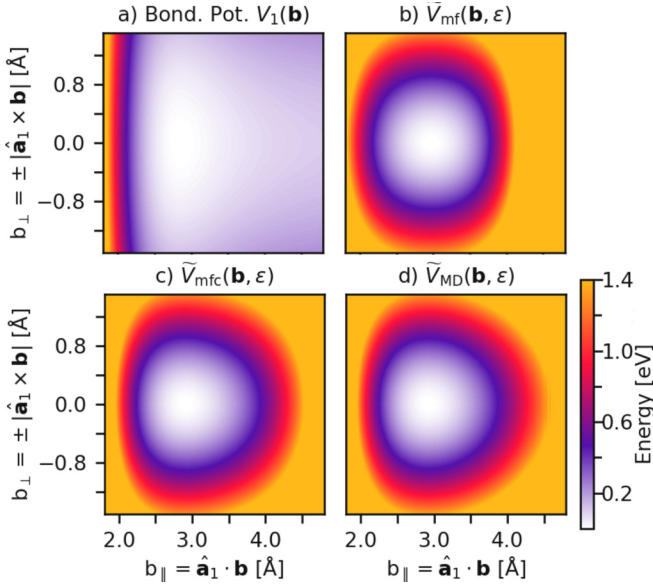
FIG. 2. Comparison of mean-field models and empirical data using (10) with $\kappa = 0.2$, $\epsilon = 0.04$, and $T = 1000$ K. (a) Two-dimensional plot of the bonding potential $V_1(\mathbf{b})$, (b) the mean-field effective potential (12), (c) the correlated mean-field effective potential (14), and (d) the empirical MD data.

1 core hour of computational effort, for each point in the (κ, T, ϵ) parameter space. The ensemble errors in $\mathcal{U}^{\mathrm{ah}}(T, \epsilon)$ and $T_{\mathrm{vir}}^{\mathrm{MD}}(T)$ were around 0.5 meV/atom and 0.5 K for a target temperature of $T = 1000$ K, which gave propagated errors in $\mathcal{F}^{\mathrm{ah}}(T, \epsilon)$ of around 1 meV/atom. This emphasizes that accurate determination of anharmonic free energies by MD sampling is numerically challenging even for simple potential energy functions such as (10), precluding the use of such models in high-throughput problems.

Mean field of a bond lattice. The key idea in our mean-field approaches is to approximate the single bond density $\rho_1(\mathbf{b}, \epsilon, T)$. To aid comparison with the empirical data we shall focus on the effective potential in which a single bond resides, given by the Boltzmann inverse of the single bond density as $\widetilde{V}(\mathbf{b}, \epsilon, T) \equiv -k_B T \ln |\rho_1(\mathbf{b}, \epsilon, T)|$. As can be seen in Figs. 2(a) and 2(d), the potential $V_1(\mathbf{b})$ of a single isolated bond is qualitatively different from the empirical effective potential $\widetilde{V}_{\mathrm{MD}}(\mathbf{b}, \epsilon, T)$, demonstrating the need to take bond-bond interactions into account. We first imposed the following mean-field approximation to (3):

$$\mathbf{b}_i^{(l)} - \mathbf{b}_i^{(k)} = \langle \mathbf{b}^{(l)} \rangle - \langle \mathbf{b}^{(k)} \rangle = \mathbf{a}_l^{\epsilon} - \mathbf{a}_k^{\epsilon}. \qquad (11)$$

It is simple to show that this condition satisfies the connectivity constraint $\mathcal{C}(\{\mathbf{b}\}) = 0$ by construction. The physical picture behind (11) is that two neighboring bonds do not move independently. Rather, by selecting one bond for the estimation of $\rho_1(\mathbf{b}, \epsilon, T)$ under (11) the neighboring bonds are fully determined. The effective potential therefore has contributions from the selected bond and all neighboring bonds. In practice, implementation of (11) in the single bond density (6) gives a
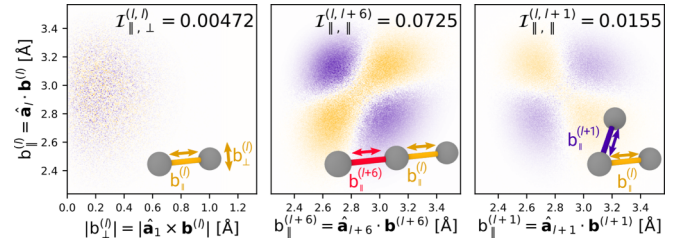

FIG. 3. Density difference $\rho_2(b, b') - \rho_1(b)\rho_1(b')$ for fcc bond component pairs $(b_\|^{(l)}, b_\perp^{(l)})$, $(b_\|^{(l)}, b_\|^{(l+6)})$, and $(b_\|^{(l)}, b_\|^{(l+1)})$ at $T = 500$ K, $\epsilon = 0$, and $\kappa = 0.1$. The degree of correlation is given by the mutual information $\mathcal{I}$ [19] of $\rho_2(b, b')$ from $\rho_1(b)\rho_1(b')$ (Supplemental Material [20]). The color scale is uniform, with orange positive and blue negative. The antiparallel $(l, l+6)$ bond pair has noticeably (approximately five times) stronger correlation.

general effective potential of

$$\widetilde{V}_{\mathrm{mf}}(\mathbf{b}, \epsilon) = \frac{1}{2} \sum_l^m V_l(\mathbf{b} + \mathbf{a}_l^{\epsilon} - \mathbf{a}_1^{\epsilon}), \qquad (12)$$

which for the single bonding potential of the fcc lattice reduces to $\widetilde{V}_{\mathrm{mf}}(\mathbf{b}, \epsilon) = \sum_{l=1}^{12} V_1[\mathbf{G}_{1l}(\mathbf{b} - \mathbf{a}_1^{\epsilon}) + \mathbf{a}_1^{\epsilon}]/2$. The effective potential, shown in Fig. 2(b), gives dramatically improved agreement with $\widetilde{V}_{\mathrm{MD}}(\mathbf{b}, \epsilon, T)$. As the fcc lattice is centrosymmetric, $V_{\mathrm{mf}}(\mathbf{b}, \epsilon, T)$ has reflection symmetry in $\mathbf{b} - \mathbf{a}_1^{\epsilon}$ and thus $\langle \mathbf{b} \rangle = \mathbf{a}_1^{\epsilon}$, meaning the mean-field model predicts the correct system volume. We emphasize that in contrast to the core hour of effort required to produce $\widetilde{V}_{\mathrm{MD}}(\mathbf{b})$ for a given parameter choice, evaluation of (12) is effectively instantaneous. An approximate single bond density can be readily obtained through $\rho_1^{\mathrm{mf}}(\mathbf{b}, \epsilon, T) \equiv \mathcal{N} \exp[-\widetilde{V}_{\mathrm{mf}}(\mathbf{b}, \epsilon)/(k_B T)]$. Evaluating the expectations (5) and (7) yields estimations $\mathcal{U}_{\mathrm{mf}}^{\mathrm{ah}}(T, \epsilon)$ and thus $\mathcal{F}_{\mathrm{mf}}^{\mathrm{ah}}(T, \epsilon)$, as shown in Fig. 4. While the correct qualitative trend is observed, this initial mean-field model underestimates the anharmonic free energy by around a factor of 2, breaching our target meV/atom error. The reason for this discrepancy is that the compatibility condition defined by Eq. (11) is too strict, forcing the mean-field model to inhabit too small a region of the full bond space defined by Eq. (3). Specifically, the condition that the difference between two neighboring bond vectors has to be a lattice vector restricts the system in finding the energetically most favorable configurations.

To identify which of the many bond-bond constraints given by Eq. (8) are not sufficiently well described we used our MD simulations to investigate correlations between various components of the bond vectors that go beyond the mean-field approximation. As illustrated in Fig. 3, this analysis revealed the perpendicular and parallel components $b_\|^{(l)}, b_\perp^{(l)}$ are essentially uncorrelated, while the antiparallel bond pairs $(l, l+6)$ (Fig. 1) show significantly larger correlation than all other bond pairs as measured by the mutual information [19] (Supplemental Material [20]). Configurations where these bonds are both longer/shorter than the average occur with enhanced probability (manifest as a depleted density along the main diagonal in Fig. 3). This behavior is reminiscent for acoustic phonons, which are expected to dominate in thermodynamic sampling due to their lower energy. It is, however, opposite
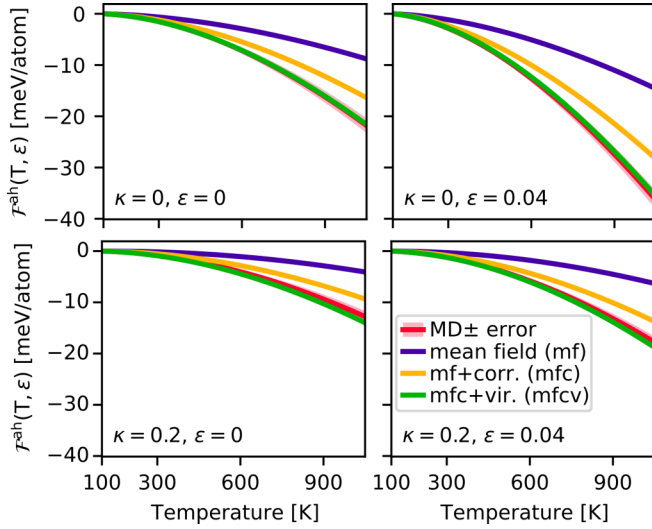
FIG. 4. The anharmonic contribution to the free energy $\mathcal{F}^{\mathrm{ah}}(T, \epsilon)$ from MD simulation, the uncorrelated mean-field model (mf), the correlated mean-field model (mfc), and the virial corrected mean-field model (mfcv) for $\epsilon = 0, 0.04$ and $\kappa = 0, 0.2$. The mfcv model is less than 1.2 meV/atom from the empirical data in all cases, within the sampling error.

to our mean-field approach according to which the sum of two neighboring bond vectors is constant and thus shows anticorrelation, i.e., if one bond is longer the other should be shorter. To retain the dominant correlation effect along chains of antiparallel bonds, we take a "correlated" mean-field approximation

$$\mathbf{b}_i^{(l)} - \mathbf{b}_i^{(k)} = \mathbf{a}_l^\epsilon - \mathbf{a}_k^\epsilon, \quad k \neq l + 6, \tag{13}$$

i.e., we assume all bonds apart from the antiparallel pairs are independent. Implementation of (13) reduces the general expression (6) for $\rho_1(\mathbf{b}, \epsilon, T)$ to a product of integrals over independent linear chains of bonds. As there is only one independent bond per atom for this quasi-one-dimensional chain, any set of bond vectors satisfying the finite volume constraint is a valid configuration, meaning $\mathcal{C}(\{\mathbf{b}\}) = \sum_{i=1}^{N_1} \mathbf{b}_i - N_1 \mathbf{a}_l^\epsilon$. In the Supplemental Material [20] we evaluate the single bond density analytically, giving a new effective potential

$$\widetilde{V}_{\mathrm{mfc}}(\mathbf{b}, \epsilon, T) = \widetilde{V}_{\mathrm{mf}}(\mathbf{b}, \epsilon) + \lambda(\epsilon, T)\hat{\mathbf{a}}_1 \cdot \mathbf{b}$$
$$+ \tfrac{1}{2}\big[V_1(\mathbf{b}) - V_7\big(\mathbf{b} - 2\mathbf{a}_1^\epsilon\big)\big], \tag{14}$$

where the scalar $\lambda(\epsilon, T)$ is determined such that $\langle \mathbf{b} \rangle = \mathbf{a}_1^\epsilon$ at a temperature T. This is achieved by minimizing a suitable objective function [15]; in our tests this required only tiny fractions of a core second on a modern processor using SCIPY.MINIMIZE [21]. As can be seen in Fig. 2(c), (14) yields an effective potential that is visually almost indistinguishable to the empirical density. Taking averages over the resultant single bond distribution

$$\rho_1^{\mathrm{mfc}}(\mathbf{b}, \epsilon, T) \equiv \mathcal{N} \exp[-\widetilde{V}_{\mathrm{mfc}}(\mathbf{b}, \epsilon, T)/(k_B T)], \tag{15}$$

gives a much improved estimation of $\mathcal{F}^{\mathrm{ah}}$, shown in Fig. 4, with a typical error of 4–5 meV/atom at 1000 K.

As a final step to reach meV/atom accuracy, we investigated thermodynamic properties of the correlated mean-field

approximation (11). We confirmed the equipartition relation $\lim_{T \to 0} \mathcal{U}_{\mathrm{mfc}}^{\mathrm{ah}}(T, \epsilon) = O(T^2)$, which indeed is required for the anharmonic component of the free energy (7) to be well defined. To ensure the mean-field model predicts the correct bond lattice temperature we used the exact expression for the virial temperature of an fcc bond lattice

$$T_{\mathrm{vir}}(T) \equiv (2/k_B) \int \mathbf{b} \cdot \nabla V_1(\mathbf{b}) \rho_1(\mathbf{b}, \epsilon, T) d\mathbf{b}, \tag{16}$$

which is derived in the Supplemental Material [20]. Substituting ρ_1^{mfc} for ρ_1 gives the mean-field virial temperature

$$T_{\mathrm{vir}}^{\mathrm{mfc}}(T) \equiv \frac{2}{k_B} \int \mathbf{b} \cdot \nabla V_1(\mathbf{b}) \rho_1^{\mathrm{mfc}}(\mathbf{b}, \epsilon, T) d\mathbf{b}. \tag{17}$$

The exact virial temperature agrees with the target temperature, i.e., $T_{\mathrm{vir}}(T) = T$, as expected. However, while $T_{\mathrm{vir}}^{\mathrm{mfc}}(T) \to T$ as $T \to 0$ K, at high temperature $T_{\mathrm{vir}}^{\mathrm{mfc}}(T)$ was found to slightly underestimate T by up to 4%. To correct for this deviation, we numerically evaluated the effective temperature $\widetilde{T}$ such that $T_{\mathrm{vir}}^{\mathrm{mfc}}(\widetilde{T}) = T$. As the Lagrange multiplier $\lambda(\epsilon, T)$ in the effective potential (14) is also temperature dependent this procedure took around a core second for each value in the (T, κ, ϵ) parameter space studied, remaining nearly four orders of magnitude more efficient than brute force sampling using the empirical bonding potential $V_1(\mathbf{b})$ in (8). We then build a new "virial corrected" single bond density using the effective temperature

$$\rho_1^{\mathrm{mfcv}}(\mathbf{b}, \epsilon, T) \equiv \rho_1^{\mathrm{mfc}}(\mathbf{b}, \epsilon, \widetilde{T}), \quad T_{\mathrm{vir}}^{\mathrm{mfc}}(\widetilde{T}) \equiv T. \tag{18}$$

Averages over $\rho_1^{\mathrm{mfcv}}(\mathbf{b}, \epsilon, T)$ thus enforce not just material compatibility but consistent thermodynamics by ensuring the true temperature of the mean-field model, namely, the predicted virial temperature of the bond lattice, matches the desired target temperature. As can be seen in Fig. 4, this gives almost perfect reproduction of $\mathcal{U}^{\mathrm{ah}}$ and $\mathcal{F}^{\mathrm{ah}}$, with less than 1.2 meV/atom error across the entire parameter space, within the ensemble error of brute force thermodynamic integration.

The very light computational demand of our mean-field model compared to brute force sampling with (8) can be verified by running the analysis scripts online [16]. We note that an equivalent sampling effort with *ab initio* molecular dynamics requires on the order of 1 core hour of computation *per timestep*, which has motivated the development of advanced thermodynamic integration strategies to reduce sampling error [10].

Conclusions. In this Rapid Communication we have provided an analytic mean-field bond model which achieves meV/atom accuracy to the anharmonic free energy of fcc bond lattices with no sampling and negligible computational cost. When suitably parametrized, bond lattices are able to reproduce *ab initio* free energies obtained by thermodynamic integration to within meV/atom [10], meaning (14) is a highly efficient and accurate surrogate model for anharmonic thermodynamic properties.

We have focused on unary fcc lattices, where a single bonding potential is sufficient to accurately reproduce *ab initio* vibrational free energies. In future work we will generalize our approach to a wider range of crystal structures, where it is expected that multiple bonding potentials will

be required. However, given the importance of fcc materials to materials science, the efficiencies demonstrated by the present model provides many opportunities for computational materials modeling, beyond the acceleration of *ab initio* free-energy calculations. Parametrization of the bonding potential requires around 10–100 force calls [10], which is essentially instantaneous even for the most advanced empirical interatomic potentials [22,23]. Using the developed model to efficiently and accurately predict the anharmonic free energy for a given potential parametrization has clear advantages for computational materials science problems where many such evaluations are needed, such as high-throughput crystal structure prediction [12]. In addition, the efficient connection between potential parameters and the anharmonic free energy could be used in the objective function during potential parametrization, to target agreement with high temperature *ab initio* phase free energies. Finally, we speculate that the structurally and thermodynamically accurate mean-field bond density (18) may allow classical density functional theory [24,25] or phase field crystal models [26,27] to more accurately represent the configuration space of atomic vibrations at high temperatures.

[1] M. Pozzo, C. Davies, D. Gubbins, and D. Alfe, Nature (London) **485**, 355 (2012).

[2] G. Grimvall, B. Magyari-Köpe, V. Ozoliņš, and K. A. Persson, Rev. Mod. Phys. **84**, 945 (2012).

[3] B. Grabowski, L. Ismer, T. Hickel, and J. Neugebauer, Phys. Rev. B **79**, 134106 (2009).

[4] T. Lelièvre, G. Stoltz, and M. Rousset, *Free Energy Computations: A Mathematical perspective* (Imperial College Press, London, 2010).

[5] O. Hellman, I. A. Abrikosov, and S. I. Simak, Phys. Rev. B **84**, 180301(R) (2011).

[6] O. Hellman and I. A. Abrikosov, Phys. Rev. B **88**, 144301 (2013).

[7] I. Errea, M. Calandra, and F. Mauri, Phys. Rev. B **89**, 064302 (2014).

[8] O. Hellman, P. Steneteg, I. A. Abrikosov, and S. I. Simak, Phys. Rev. B **87**, 104111 (2013).

[9] A. Glensk, B. Grabowski, T. Hickel, and J. Neugebauer, Phys. Rev. X **4**, 011018 (2014).

[10] A. Glensk, B. Grabowski, T. Hickel, and J. Neugebauer, Phys. Rev. Lett. **114**, 195901 (2015).

[11] S. G. Moustafa, A. J. Schultz, E. Zurek, and D. A. Kofke, Phys. Rev. B **96**, 014117 (2017).

[12] R. Bianco, I. Errea, L. Paulatto, M. Calandra, and F. Mauri, Phys. Rev. B **96**, 014111 (2017).

[13] T. D. Swinburne and M.-C. Marinica, Phys. Rev. Lett. **120**, 135503 (2018).

[14] A. S. Bochkarev, A. van Roekeghem, S. Mossa, and N. Mingo, Phys. Rev. Materials **3**, 093803 (2019).

[15] T. Swinburne and J. Janssen, BLaSA code, https://github.com/tomswinburne/BLaSA.

[16] T. Swinburne and J. Janssen, Binder-BLaSA analysis, https://mybinder.org/v2/gh/tomswinburne/BLaSA/master?filepath=analyze.

[17] Equation (3) can be derived in the continuum limit using Stoke's law [28] for the tensor field $yy \oplus_{l=1}^{d} \mathbf{b}^{(l)}(\mathbf{R}_i)$.

[18] Accounting for the $d(m/2 - d)N$ additional constraints when $m > 2d$, we have $mdN/2 - d(d - 1)N - d(m/2 - d)N = dN$ degrees of freedom as required.

[19] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, New York, 2012).

[20] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevB.102.100101 for a derivation of the mean-field effective potential, the bond lattice virial temperature, and further details on the bond correlation analysis.

[21] P. Virtanen *et al.*, Nat. Methods **17**, 261 (2020).

[22] A. M. Goryaeva, J.-B. Maillet, and M.-C. Marinica, Comput. Mater. Sci. **166**, 200 (2019).

[23] Y. Zuo, C. Chen, X. Li, Z. Deng, Y. Chen, J. Behler, G. Csányi, A. V. Shapeev, A. P. Thompson, M. A. Wood *et al.*, J. Phys. Chem. A **124**, 731 (2020).

[24] K. R. Elder, N. Provatas, J. Berry, P. Stefanovic, and M. Grant, Phys. Rev. B **75**, 064107 (2007).

[25] J. F. Lutsko, Adv. Chem. Phys. **144**, 1 (2010).

[26] A. Jaatinen, C. V. Achim, K. R. Elder, and T. Ala-Nissila, Phys. Rev. E **80**, 031602 (2009).

[27] G. Demange, H. Zapolsky, R. Patte, and M. Brunel, npj Comput. Mater. **3**, 1 (2017).

[28] R. Fosdick and G. Royer-Carfagni, Int. J. Non-Linear Mech. **40**, 381 (2005).