


Machine learning for metallurgy I. A neural-network potential for Al-Cu

Daniel Marchand , Abhinav Jain , Albert Glensk , and W. A. Curtin

Institute of Mechanical Engineering, École Polytechnique Fédérale de Lausanne, CH-1015 Vaud, Switzerland

 (Received 24 June 2020; revised 18 August 2020; accepted 9 September 2020; published 1 October 2020)

High-strength metal alloys achieve their performance via careful control of precipitates and solutes. The nucleation, growth, and kinetics of precipitation, and the resulting mechanical properties, are inherently atomic scale phenomena, particularly during early-stage nucleation and growth. Atomistic modeling using interatomic potentials is a desirable tool for understanding the detailed phenomena involved in precipitation and strengthening, which requires length and timescales far larger than those accessible by first-principles methods. Current interatomic potentials for alloys are not, however, sufficiently accurate for such studies. Here a family of neural-network potentials (NNPs) for the Al-Cu system are presented as a first example of a machine learning potential that can achieve near-first-principles accuracy for many different metallurgically important aspects of this alloy. High-fidelity predictions of intermetallic compounds, elastic constants, dilute solid-solution energetics, precipitate-matrix interfaces, generalized stacking fault energies and surfaces for slip in matrix and precipitates, antisite defect energies, and other quantities, are shown. The NNPs also captures the subtle entropically induced transition between θ' and θ at temperatures around 600 K. Many comparisons are made with the state-of-the-art angular-dependent potential for Al-Cu, demonstrating the significant quantitative benefit of a machine learning approach. A preliminary kinetic Monte Carlo study shows the NNP to predict the emergence of GP zones in Al-4at%Cu at $T = 300$ K in agreement with experiments. These studies show that the NNP has significant transferability to defects and properties outside the structures used to train the NNP but also shows some errors highlighting that the use of any interatomic potential requires careful validation in application to specific metallurgical problems of interest.

DOI: [10.1103/PhysRevMaterials.4.103601](https://doi.org/10.1103/PhysRevMaterials.4.103601)

I. INTRODUCTION

The mechanical properties of metals are related to the underlying behavior of defects in the crystalline lattice. The atomic structures of dislocations, grain boundaries, interfaces, precipitates, crack tips, and vacancies and their motion, evolution, and/or interactions all determine the plastic flow, fracture toughness, creep, fatigue, radiation resistance, and other essential macroscopic performance measures. Understanding, and ultimately controlling, the behavior of these defects is crucial for optimizing application conditions and designing new higher-performance alloys. This, in turn, requires atomic-scale simulations, but the relevant structures are often too large to be accessible by first-principles methods, such as density-functional theory (DFT). One solution is thus the development of semiempirical interatomic potentials that accurately capture the structures, energies, and motion of the various defects.

Potentials for metals, mainly within the embedded-atom or modified embedded-atom (EAM and MEAM, respectively) frameworks, are widespread and generally perform well against experimental or first-principles benchmarks to which they are fit. However, potentials can be very inaccurate for certain properties, e.g., unstable stacking fault energies, or can predict unphysical behavior for a given defect, e.g., a sharp crack tip. Potentials for alloys generally have even more limited capabilities due to the underlying assumptions

of the embedded atom framework. Alloy potentials are thus only reasonably quantitative for a few properties and are not accurate enough for complex defects and defect interactions. The shortcomings of existing potentials stem in part from the relatively rigid functional forms of the EAM and MEAM frameworks, especially limiting for multielement interactions except in a few rare cases (cf. Ref. [1]). Hence, a different approach is needed so that the power of atomistic computations can be used to understand, predict, and design the performance of technologically valuable metal alloys.

Machine learning (ML) offers a new approach to fit the potential energy surface (PES) of a metal without imposing a highly restricted functional form [2–6]. The construction of an ML potential consists of (i) developing a training set or database of energies and forces of atomic structures, (ii) choosing suitable descriptors for the local atomic environments across the training dataset [7–10], and (iii) applying a selected regression algorithm (e.g., neural network, kernel ridge regression) to optimize the parameters in the ML framework so as to best match the training data. Since the number of descriptors and parameters is unlimited in principle, the ML approaches provide a “parameter-rich” space to fit the many configurations in the PES to the training dataset. However, ML potentials usually fail to extrapolate to regions of phase space outside the training set. Furthermore, the choices of atomic descriptors, ML framework, and extent of the training data all affect the final potential.

It should be recognized, however, that the above caveats regarding ML potentials are present with traditional interatomic potentials as well. That is, the fixed functional form of an EAM or MEAM potential with limited parameters intrinsically limits the ability to fit many properties accurately. Therefore, they still require user-imposed decisions about which properties are most significant and which are not. The perceived advantages of the traditional potentials are that (i) the concept of embedding is well founded in the underlying quantum mechanics of the problem, and (ii) specific desired behavior (e.g., very strong short-range atomic repulsions) can be easily enforced at the outset. However, traditional potentials also have many hidden parameters, e.g., the complex spline fits of the “electron density” of an atom and the arbitrary form of the added pair potential, that are used to best-fit the training data. The choice of traditional (e.g., EAM or MEAM) versus ML potentials can be reduced to a choice between either:

(i) EAM or MEAM, where the potential is built from physically derived smooth functions and so less likely to have egregious extrapolation errors but is still highly fitted and unable to capture many necessary materials properties

or

(ii) ML, where the potential is purely a regression on the training set and, being parameter-rich, can fit many desired properties but may exhibit unphysical behaviors in regions of extrapolation.

To minimize issues with ML potentials, curation of the training dataset is essential. Ideally, an exhaustive dataset of possible atomic environments with their energies and forces from electronic structure calculations is needed to ensure transferability. Generating the training dataset starts with defining the physical properties that the potential should reproduce and the associated atomic configurations. Representative atomic environments are then selected for the actual training dataset (see Fe [11] and W [12] within the GAP framework, for instance). Protocols to circumvent the tedious task of curating an exhaustive training dataset have been developed. One method, *active learning*, has used random perturbations of bulk crystalline structures to sample configuration space in an Al-Mg alloy [13]. A second method, *self-guided learning*, has explored the phase space using randomized unit cells paired with a selection diverse structures with application to C, Si, and Ti [14]. A third method, *on-the-fly learning*, has combined density-functional-theory (DFT) calculations with ML to evaluate melting points for Al, Si, Ge, Sn, and MgO [15]. This approach circumvents expensive DFT calculations of 99% of structures because only structures with an estimated error larger than a defined bound are recalculated via DFT and then added to improve the ML force field. A hybrid approach, called the physically informed neural-network (PINN) method, combines an analytical bond order potential (BOP) with a neural network that adjusts the BOP parametrization as a function of the specific environment [5]. Such a potential retains some physical bounds when extrapolated to new atomic environments.

A limiting aspect of nearly all of the ML potentials generated to date is that the training data, and fitness of the potential, are mainly demonstrated on fundamental properties

of the bulk crystalline material. Few material defects are considered. For example, there has been extensive work in the literature to create ML models to capture the entire periodic table [16–18]; however, these studies were restricted to bulk configurations and typically have absolute energy errors of 70–80 meV/atom, much higher than the 10 meV/atom typical of a Behler-Parinello (B-P) potential [3]. Structures in the training set are associated with the equilibrium geometry, elastic response of the bulk, vibrational modes, single vacancies, selected surfaces, and sometimes liquid-state information. These features are absolutely necessary but far from sufficient for performing metallurgically useful studies of the behavior of defects in metals.

A few efforts have extended beyond basic properties. The GAP potentials for Fe [11] and W [12] included baseline data needed for describing dislocations, and the GAP Fe potential was used to study the double-kink nucleation process that controls plastic flow in BCC metals [19]. Kobayashi *et al.* [4] developed a neural-network potential (NNP) for the ternary Al-Mg-Si system including a wide range of intermetallics, solute-solute interactions, and interfaces and showed good predictions for edge and screw dislocation structures, solute-dislocation interactions, and *in situ* precipitates; further development was shown by Imbalzano *et al.* [20]. The PINN method [5] showed application to an FCC edge dislocation. Many of these properties are difficult to compute with *ab initio* methods and are still only the next step toward realistic metallurgical studies (e.g., dislocations interacting with precipitates or fracture in an alloy).

In this work, we focus on Al-Cu as a case study for atomic modeling of alloys with a ML potential. The Al-Cu alloys are well studied [21] and are crucial to the automotive and aerospace industries [22]. Al alloy production is one of the great energy consumers of the planet [23], providing further motivations for both enhancing performance and understanding chemistry to facilitate recycling. A good Al-Cu potential needs to capture many structures and energies: e.g., solute clusters, bulk and interface stability of precipitates, generalized stacking fault energies (GSFEs) of precipitates and matrix, and vacancy and antisite formation energies. Furthermore, Al-Cu has a subtle entropically driven $\theta' \rightarrow \theta$ transformation [24] that should be reproduced by any suitable potential.

Apostol and Mishin created an angular-dependent potential (ADP) for Al-Cu [25], which is an extension of the EAM formalism to include angular terms. The potential was designed to predict certain formation energies, lattice constants, elastic constants, surface energies, and generalized stacking fault energies. This potential has now been used in numerous studies of mechanical behavior of precipitates and in studying solute and precipitate strengthening [26–28]. In spite of these applications, the Al-Cu ADP has not been examined across a full range of properties or configurations necessary for broad metallurgical studies. Given the general acceptance of EAM and MEAM potentials and the high concern about ML potentials as being purely fitted without underlying physics, the Al-Cu ADP serves as an excellent benchmark against which to compare any new ML potential in this alloy. Here, while we will show many deficiencies of the Al-Cu ADP potential,

this is not done as a means of discrediting this or any other traditional potential. Rather, our point is to highlight the fact that even the best traditional potentials for alloys perform far worse than an ML potential that is developed using a broad training dataset. The general rule is that all potentials, traditional or ML, must be carefully and widely validated to the best level possible prior to their application to any metallurgical problem.

With the above background, our goal in this work is construct a neural-network machine learning potential (NNP) for the Al-Cu system that can set the state of the art for metallurgical modeling in a metal alloy. We carefully and systematically examine many atomic-level properties relevant, with a particular emphasis on solutes and precipitates. We then show that the NNP is broadly successful in capturing defects and metallurgical phenomena that are not within the training dataset and that this performance is far better than that of the ADP potential. While executed here for only one binary alloy, we believe this example provides a significant driving force for the development of such machine learning potentials for metal alloys more broadly. Such new ML potentials will then greatly enhance the quantitative quality of atomistic modeling of alloys, enabling better mechanistic understanding of performance-controlling atomistic processes and so helping to accelerate improvements to existing and emerging alloys.

The remainder of this paper is structured as follows. In Sec. II, we present the methodology and details of the metallurgical dataset. In Sec. III, we assess the ability of the NNPs to represent the training dataset and properties associated with data derived from the dataset. In Sec. IV, the NNPs are then examined with respect to many metallurgical issues, including precipitation, interfaces, dislocations and stacking faults, and fracture. In Sec. V we summarize our findings and provide some perspective.

II. METHODOLOGY

In this section, we describe all of the atomic structures used here for modeling metallurgical phenomena in Al-Cu. The associated energies and forces as computed using density-functional theory serve as a database for the application of any machine learning method to develop an interatomic potential. The details of the first-principles DFT methodology are then presented, including our development of general workflows within the framework of the AiiDA infrastructure [29]. We then describe the neural-network approach adopted here, which is that proposed by Behler and Parinello [7].

A. Atomic structures

We wish to develop a comprehensive set of atomic structures to train a machine learning potential that is broadly useful for modeling precipitation, plasticity, and fracture behavior in Al-Cu alloys. While we include structures related to pure Cu and Al in Cu, our main goal is a potential for Al-Cu in the range of dilute Cu that is relevant for technological applications of these alloys. To this end, we created 4857 structures within the following categories:

(i) All Al-Cu intermetallic structures from the Open Quantum Material Database (OQMD) [30]. θ and θ'' structures were also added.

(ii) (111), (110), and (100) surfaces for pure Al and Cu.

(iii) Stable and unstable stacking faults for pure Al and Cu.

(iv) Additional Al, Cu, θ , θ' , and θ'' structures at high pressures.

(v) Binary solute-solute and solute-vacancy pairs at varying distances in both Al and Cu matrices.

(vi) Solutes at stacking faults in Al and Cu.

(v) Thirty-two atom FCC supercells with random concentrations of Al, Cu, and vacancies.

(vi) Coherent and semicoherent interfaces between Al and θ'' and θ' precipitates,

(vii) Generalized stacking fault (GSF) ($0\bar{1}1$) surface for the θ precipitate.

(viii) Supercells of all OQMD structures with small displacements suitable for phonon calculations

Specifically, we fully relaxed the cell shape and atomic positions of 60 OQMD structures plus θ and θ'' phases using the methodology described in Sec. II B. 45 of the relaxed structures maintained the initial symmetry group and were considered “stable” (many structures in the OQMD relaxed into the same structure, e.g., pure non-FCC Al and Cu structures would often relax into the FCC phase). For each of the stable structures, further structures were generated by applying strains of magnitudes -0.01 , -0.005 , 0.005 , and 0.01 in each of the $\epsilon_{1..6}$ strain directions (Voigt notation) and reduced to a set of unique independent strains using Pymatgen [31] (e.g., for FCC the necessary independent strains are only ϵ_1 and ϵ_4). Surfaces and stacking fault structures in pure Al and Cu were generated using the Atomic Simulation Environment (ASE) software package.

Binary Cu-Cu, Cu-vac, and vac-vac pairs embedded in an Al matrix were generated up to seventh nearest-neighbor distances in a 256 atom supercell. Binary Al-Al, Al-vac, and vac-vac pairs embedded in a Cu matrix were generated up to fourth nearest-neighbor distances in a 108 atom supercell. Thirty-two atom FCC supercells spanning all Al-Cu concentrations ratios in increments of 10%, as well as vacancy concentrations of 5% and 10%, were generated for a total of 990 structures (30 for each concentration). For each generated structure the lattice constant was fixed to be a concentration-weighted average between Al (4.07 Å) and Cu (3.68 Å) where the large Al lattice constant was selected to better-simulate high temperatures. For each atomic position the occupying atom was selected randomly according to a probability equal to the concentration and was also shifted from its lattice site using by a small random vector with amplitude selected from a Gaussian probability distribution with a standard deviation 0.15 Å.

Solute or stacking fault structures in Al and Cu were generated by first creating a stacking fault structure and then placing a single solute at an atomic position in the first atomic plane adjacent to the fault plane. Four Al-matrix-precipitate interfaces structures were created: semicoherent θ'' , coherent θ'' , semicoherent θ' , and coherent θ' . The θ phase does not have a coherent interface [22] with Al-bulk and so was not studied. We used three different sizes for the θ' interfaces

and four different sizes for the θ'' interfaces, with each size corresponding to an equal number of layers of precipitate and matrix on either side of the interface. Structures corresponding to high-symmetry points of the $(0\bar{1}1)\theta$ generalized stacking fault surface were also created (this structure having also some particular importance for laser-sintered laminated structures [6]). Phonons in the 45 structurally stable OQMD structures were studied using Phonopy [32] with supercells having a minimum of 108 atoms and with cell shapes as close to cubic as possible. Phonopy was then used to determine the dynamical matrix; e.g., a 144 atom supercell for the $\text{Al}_2\text{Cu}-\theta$ phase requires three displacements (two on an Al atom and one on a Cu atom).

To test the NNPs developed using the above structures, we also examined a number of other structures or properties that were not included in the above training set

- (i) GSF for the (111) plane of θ''
- (ii) Cu clusters of three, four, and more atoms in an Al matrix
- (iii) Unrelaxed antisite and vacancy energies for all sites in all OQMD structures
- (iv) Thermodynamics of the θ/θ'' transition
- (v) Melting of Al and eutectic Al-18%Cu
- (vi) Early precipitate clusters, i.e. Guinier-Preston (GP) zone formation during annealing as predicted using kinetic Monte Carlo (KMC); there are no first-principles results for these structures
- (vii) The $a\langle 110 \rangle/2$ edge dislocations in pure Al
- (viii) An atomistically sharp crack and dislocation emission under load; there are no precise first-principles results for these structures.

While amorphous Al-Cu has been studied recently [33], it is not a major technological alloy and atomistic models have difficulty reproducing amorphous structures consistent with realistic cooling rate and so we do not study it here.

B. DFT methodology

All first-principles calculations were performed within the framework of DFT as implemented using Quantum Espresso [34] (QE). The GGA-PBE [35] exchange-correlation functional was used with a 544-eV (40-Ry) energy cutoff. An 80 k -points/ \AA^{-1} Monkhorst-Pack grid [36] (corresponding to a $20 \times 20 \times 20$ k -point grid for the 4.04 \AA^3 conventional cubic FCC cell) and an 0.6-eV (0.0441-Ry) Methfessel-Paxton smearing parameter [37] were also used. These latter choices were made after significant convergence and validation studies for achieving reliable elastic constants, especially C_{44} , in both Al and Cu. The Al pseudopotential was chosen from the solid-state pseudopotential (SSSP) library [38]. The Cu pseudopotential of Dal Corso [39] was used, which was found to have better accuracy and reduced computational cost as compared to the SSSP version.

The AiiDA platform [29,40] was used to manage the large number of DFT computations needed for this study, as well as to ensure that all calculations used consistent settings, which is essential for ML potentials [11]. Figure 1 shows a schematic of how AiiDA was used, with the various key terms having the following meanings:

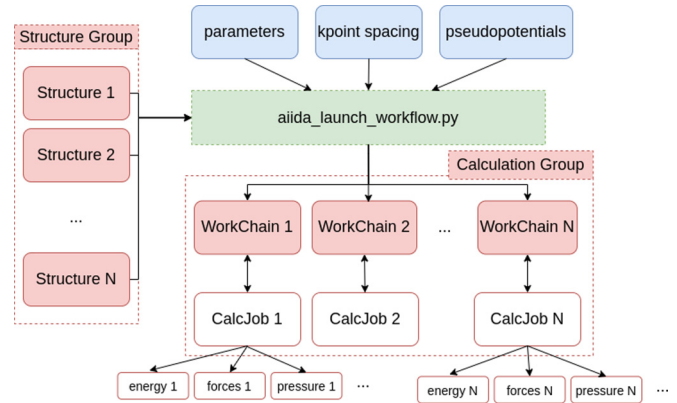


FIG. 1. Schematic of AiiDA usage. Structures are stored in the database in a Structure Group; jobs are launched with the `aiida_launch_workflow.py` wrapper script; each WorkChain object manages a CalcJob where common events, e.g., job submission, continuation of jobs past cluster time outs are handled automatically; outputs such as computed energies, forces, pressures, and atomic structures are stored automatically in a database, organized in Calculation Groups with typically one for each Structure Group of interest.

(i) Structure Group: A group of structures stored in the AiiDA database

(ii) parameters: A set of parameters stored in the AiiDA database. These are the settings, e.g., energy cutoff, smearing, to be used for each calculation

(iii) pseudopotentials: A set of pseudopotentials stored in the AiiDA database

(iv) kpoint spacing: The k -point distance to be used in each calculation.

(v) CalcJob: An AiiDA object that directly interfaces with an external code, in this work QE, to submit, launch, run, collect, and parse on a HPC cluster. In this work all CalcJobs are created and managed by a parent WorkChain.

(vi) WorkChain: An AiiDA object that encapsulates a set of CalcJobs to achieve a high-level workflow goal. For example, a cell relaxation requiring multiple restarts or to compute elastic constants of a material.

(vii) `aiida_launch_workflow.py`: A wrapper script for launching WorkChains

(viii) Calculation Group: A group of WorkChains in the AiiDA database

Because AiiDA can serve as a future platform for calculations on other alloy systems, we further detail our usage of it as follows. First, we created a set of user-friendly scripts within AiiDA to create each type of structure mentioned in Sec. II A. These structures were stored in AiiDA Structure Groups. In some cases, e.g., for importing the OQMD structures or generating phonon-displaced structures, we used an external script to create structures, which we then uploaded into AiiDA. A wrapper script enabled the application of a consistent set of parameters, k -point spacing, and pseudopotentials to generate an AiiDA WorkChain for each structure in the Structure Group. Each WorkChain managed a QuantumEspresso calculation as a CalcJob on an HPC cluster, where events such as job time outs and result parsing were

handled automatically. The `aaida_launch_workflow.py` wrapper script could launch in several modes to create atomic-relaxation, atomic and cell relaxation, or no-relaxation (SCF), as needed in each case. AiiDA stored the results of each calculation in a Calculation Group. These files were later parsed using a dedicated Jupyter [41] notebook for each property of interest, e.g., surface energies, GSF, elastic constants. We provide all scripts and all computed data openly on the Materials Cloud [42,43].

C. NNP selection and evaluation of errors

In this work, we generated NNPs using the methodology developed by Behler and Parrinello [7] as implemented in the open source code n2p2 [44]. Here we sketch the essentials of the method and refer the reader to the excellent tutorial review of Behler for further details [3].

An NNP predicts the energy of an entire structure by summing energies of each atom within the structure as

$$E^{\text{structure}} = \sum_i E_i^{\text{atom}}. \quad (1)$$

Each atomically local energy, E_i^{atom} , is computed using a nested hierarchical function of weighted layers across a neural network. We use a neural-network architecture with two hidden layers and 24 nodes per layer. We did not rigorously investigate other network architecture, and other configurations would likely work as well. The functional form for the atomic energy is given explicitly as

$$E_i^{\text{atom}} = f_3 \left\{ b_1^3 + \sum_{k=1}^{24} a_{k,1}^{2,3} f_k^2 \right. \\ \left. \times \left[b_k^2 + \sum_{j=1}^{24} a_{j,k}^{1,2} f_j \left(b_j^1 + \sum_{i=1}^{64} a_{i,j}^{0,1} G_i \right) \right] \right\}, \quad (2)$$

where $a_{z,w}^{q,p}$ is the weighting from node z on layer q to node w on layer p ; similarly, b_z^q is the bias of node z on layer q . f_q is an activation function; here we use the softplus function $\ln(1 + e^x)$ for f_1 with f_2 and f_3 the identity functions.

The G_i are functions used to characterize a local atomic environment and are generally referred to as the Descriptors. In the B-P NNP framework, they are called symmetry functions. We use radial and angular symmetry functions defined as

$$G_i^{\text{radial}} = \sum_{j=1}^{N_{\text{atom}}} e^{-\eta(R_{ij}-R_c)^2} f_c(R_{ij}), \quad (3)$$

$$G_i^{\text{angular}} = 2^{1-\zeta} \sum_{j \neq i} \sum_{k \neq i,j} \left[(1 + \lambda \cos \theta_{ijk})^\zeta e^{-\eta(R_{ij}^2 + R_{ik}^2 + R_{jk}^2)} \right. \\ \left. \times f_c(R_{ij}) f_c(R_{jk}) f_c(R_{jk}) \right], \quad (4)$$

where R_c , η , ζ , and λ are predefined, or hyper-, parameters for a given symmetry function, R_{ij} is the distance between atoms i and j , and $f_c(r) = \tanh^3(1 - r/r_c)$ is a cut-off function, where r_c is a cutoff parameter for each symmetry function. From among the above general class of symmetry functions, we select a subset based on the structures in our training dataset using the CURSEL method of Imbalzano *et al.* [20]. First, we generate a dense grid hyperparameters to create a total

of 1192 unique radial and angular symmetry functions. We then performed a CUR decomposition on all 4857 structures and selected the 32 most-descriptive Al-centered symmetry functions and the 32 most descriptive Cu-centered symmetry functions for a total of 64 symmetry functions. We found that 32 symmetry functions per atom type resulted in a good balance between accuracy and computational cost. The resulting list of symmetry functions is shown in the Supplemental Materials [45].

Training of the NNP potential, i.e., determination of the weights and biases to best-match the DFT reference data, was accomplished with the n2p2 library [44,46]. We selected at random 90% of the structures for training and the remainder for testing. NNP weights were updated during the training using a fading memory Kalman filter. The training is based on minimizing an objective, or loss, function. We trained a loss function Γ^{loss} on the total energy of the i th structure in the training set, E_i , as well as the j th force for that structure $F_{i,j}$:

$$\Gamma^{\text{loss}} = \frac{1}{N^{\text{struct}}} \sum_{i=1}^{N^{\text{struct}}} \left[(E_i^{\text{NNP}} - E_i^{\text{DFT}})^2 \right. \\ \left. + \frac{\beta}{3N_i^{\text{atom}}} \sum_{j=1}^{3N_i^{\text{atom}}} (F_{i,j}^{\text{NNP}} - F_{i,j}^{\text{DFT}})^2 \right], \quad (5)$$

where the sums are performed over all training structures N^{struct} at each iteration of training. Note that for each structure there is only one energy, E_i , but $3N_i^{\text{atom}}$ forces, i.e., three forces (X, Y, Z) for each atom. β controls the relative influence of forces on the loss function, in this study $\beta = 8(\text{\AA}^2)$. Each training structure of N atoms provided $3N$ forces, but only one energy, and we wished to avoid having the forces overwhelm the energies when training the potentials. We used 2.1% of forces during each training epoch, (the n2p2 default is 2.3%), increasing the ratio beyond this number resulted in prohibitive computational cost. Note that this formulation tends to be biased toward larger structures since each atom independently contributes to the error. We use a standard 150 epochs (iterations) of training per NNP after which we found there to be only minimal gains in accuracy. LAMMPS [47] was used to compute the energies and forces of all structures using the evolving NNP during training via the n2p2 interface.

Here we generated 40 NNPs that differ in the specific 90% training structures and the initial random weights and biases. The initial training structure dominates the final NNP. Note that these 40 NNP are thus not independent potentials. However, they represent the range of potentials that are achieved using this NNP structure, symmetry functions, and set of training structures. Their predictions for various material properties thus represent the range of properties associated with these NNP details. The set of 40 NNPs can be used as a committee model to assess whether a particular new structure outside the overall training dataset may or may not be accurately represented. To present one specific set of results for many different properties, we select one NNP from among the 40 NNPs, labeled NNP11, that has the lowest error for C_{44} in Al. This choice is motivated by the importance and relative difficulty of modeling C_{44} accurately. In results below,

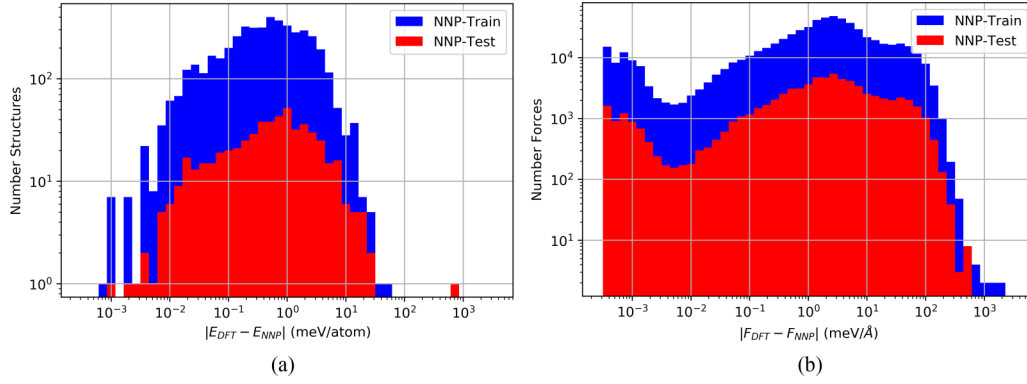


FIG. 2. Histogram of errors for the (a) energy and (b) forces for both training and testing for NNP11.

all error bars center on the average of all 40 trained NNPs and indicate ± 1 standard deviation. We note that the standard deviation is an imperfect model for the error of NNPs which we use for its extreme simplicity; for a more sophisticated treatment we refer to the work of Musil *et al.* [48]. The NNP11 may fall near the average or outside the error bars; it is just one particular NNP. This also highlights the care necessary in considering the average and standard deviation across the 40 NNPs: There may not be any single NNP that is within 1 standard deviation of the average for all properties considered. Last, we reiterate that the standard deviation does not indicate statistically independent results since all NNPs are trained on very similar sets of structures.

III. RESULTS: NNP VERSUS TRAINING AND TEST DATA

A. Overall NNP accuracy

Here we look at the root-mean-square error (RMSE) for structures in the testing and training sets for both energies and forces separately,

$$\text{RMSE}(E) = \left[\frac{1}{N_{\text{struct}}} \sum_i^{N_{\text{struct}}} (E_i^{\text{DFT}} - E_i^{\text{NNP}})^2 \right]^{\frac{1}{2}}, \quad (6)$$

$$\text{RMSE}(F) = \left[\frac{1}{N_{\text{struct}}} \sum_i^{N_{\text{struct}}} \sum_j^{3N_i^{\text{atom}}} \frac{1}{3N_i^{\text{atom}}} (F_{i,j}^{\text{DFT}} - F_{i,j}^{\text{NNP}})^2 \right]^{\frac{1}{2}}. \quad (7)$$

Over all 40 NNPs there is an average RMSE for the energy of 2.49 meV/atom for testing and 69.15 meV/atom for training, while the testing and training RMSE for the forces are 23.9 and 35.93 meV/Å, respectively. These would appear to be unacceptably large testing errors, indicating poor potential quality. In fact, most of the NNPs show only a modest difference, with 21 of the 40 potentials having a testing RMSE for energy of less than 5 meV/atom. The major contributor to the large error in the RMSE values for the remaining potentials is due to a single outlier structure under very high compression. Examining NNP11 as an example, the RMSE for energy is 2.58 meV/atom for the training set but 26.51 meV/atom for the testing set, but the histogram of errors shown in Fig. 2(a) reveals the single outlier, which is FCC Al at lattice constant $a/a_0 = 0.81$. The testing RMSE for NNP11 without this

outlier structure would fall to a very acceptable value of only 3.46 meV/atom. The forces follow a nearly uniform error distribution and do not have an outlier like in the energies, as can be seen in Fig. 2(b). For all other NNPs with high test-train error ratios, the cause was found to be outlier structures with $a/a_0 = 0.8$.

The predictions of ML potentials at high compression outside the domain of training is commonly highlighted as a failure of NNPs. In fact, such behavior is easily prevented by the addition of a strongly repulsive potential for small atomic spacings, and such an ad hoc potential is part of standard EAM or MEAM potentials to handle exactly this behavior. Of equal importance is that there are few, if any, metallurgical problems where such a high compression and high energy structure could be accessed. Here the NNPs have high error when the training set does not include the structures with $\frac{a}{a_0} = 0.8$, as can be seen in Fig. S1 in the Supplemental Material [45]. Putting aside this unusual structure—which can be captured if it is included in the training set—all 40 NNPs perform very well overall.

B. Bulk properties

Figure 3 shows the fractional deviation with respect to the DFT reference value for a number of fundamental properties in pure Al and Cu as predicted by the NNP and ADP. For the lattice and elastic constants and surface energies, to which the ADP is directly fit, the NNP shows no advantage relative to ADP. The NNP only improves over ADP in the predicted stable stacking fault energies in Al. The NNP shows substantially different values for C_{44} of Al as compared to DFT, systematically too large by 20%. Such a deviation is not unusual for NNPs, and prior work has found errors of a similar magnitude for elastic constants [49]. Elastic constants are derived from the second derivative of the energy around the minimum energy, and so are presumably challenging to model accurately when training is performed primarily on energies and on a training set with a very wide distribution of energies relative to the bulk perfect crystal. We note that the error is not in the DFT because the DFT parameters have been carefully converged with respect to C_{44} . Otherwise, for the basic properties to which EAM or MEAM-type potentials are fit, the NNP performs roughly as well as the ADP.

Figure 4 shows the formation energies, atomic volumes, and elastic constants for all Al-Cu intermetallic structures

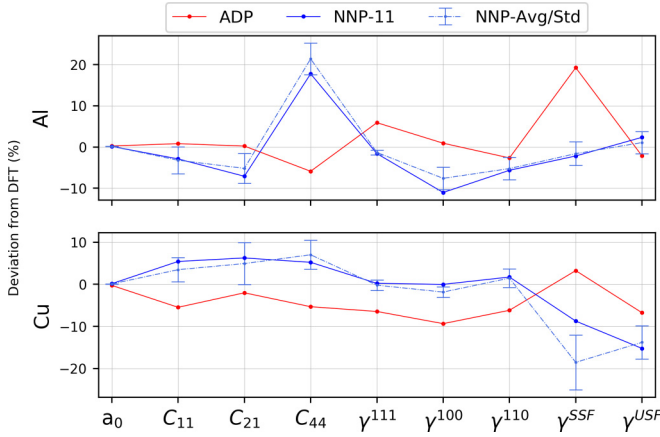


FIG. 3. Deviation of NNP and ADP properties from DFT reference values (%) for lattice constants, elastic constants, surface energies, and stable and unstable stacking fault energies of FCC Al and Cu. Dashed blue lines and error bars: mean and standard deviation across 40 NNPs; dark blue line: NNP11; red lines: ADP. All structures from which these properties are derived are included in the training set.

within the OQMD database as computed via DFT, the NNPs (mean, standard deviation, and the specific NNP11), and the ADP. The formation energy $\Delta E_f^{\text{structure}}$ of each structure is computed relative to bulk FCC Al and dilute solid-solution Cu in the Al matrix as

$$\Delta E_f^{\text{structure}} = E_N^{\text{structure}} - nE_{\text{Al}}^{\text{ref}} - (N - n)E_{\text{Cu}}^{\text{ref}}, \quad (8)$$

$$E_{\text{Al}}^{\text{ref}} = E_{\text{Al}}^{\text{FCC}}, \quad (9)$$

$$E_{\text{Cu}}^{\text{ref}} = E_{255\text{Al},1\text{Cu}}^{\text{1sol}} - 255E_{\text{Al}}^{\text{FCC}}, \quad (10)$$

where N and n are the total number of atoms and number of Al atoms, respectively, in a given structure. The formation energies are generally very well predicted by the NNPs, with most errors within a few meV/atom. In contrast, the ADP often highly overestimates or underestimates the formation energy; in a few cases, ADP reverses the sign relative to pure Al. The B2 structure poses a challenge to both ADP and most NNPs. The ADP predicts the B2 phase to be extremely stable—more stable than any other structure—while the NNP predictions have the highest error among all structures considered. Some NNPs have, however, good predictions for the B2 phase and the NNPs largely preserve the overall order of stability among the many different phases. For structures with high formation energy, both the NNP and ADP show inaccuracies, often not relaxing to the same structure as DFT or having substantial errors, especially the ADP. The ADP regularly underestimates the atomic volume, even for the critical θ and θ' structures. Since strain energies due to lattice mismatch can contribute to *in situ* precipitate formation energies, these large deviations in volume can lead to substantial erroneous effects. For elastic constants, neither the ADP nor the NNPs are extremely accurate but overall the NNPs are superior to the ADP, especially for the important low-energy structures.

C. Solutes

The binary formation energies for Cu, Al, and vacancy in Al or Cu matrix are computed as

$$\Delta E_f^{\text{bind}} = E_{N-2,X,Y}^{2\text{sol}} - E_{N-1,X}^{1\text{sol}} - E_{N-1,Y}^{1\text{sol}} + E_N^{\text{Pure}}, \quad (11)$$

where ΔE_f^{bind} is the binding energy for the solute pair, with a negative energy indicating attraction. $E_{N-2,X,Y}^{2\text{sol}}$, $E_{N-1,X}^{1\text{sol}}$, $E_{N-1,Y}^{1\text{sol}}$, and E_N^{Pure} are the energies for a solute pair, a single solute of species X, a single solute of species Y, and a system of pure bulk matrix, respectively. Here $N = 256$ for Al matrix and $N = 108$ for Cu matrix. The DFT, NNP, and ADP results for these quantities in Al are shown in Fig. 5 while those for the Cu matrix are shown in Fig. S3 in the Supplemental Material [45]. The NNP predictions for pair binding energies are generally in good agreement with DFT, typically within 25 meV (kT at room temperature) across all NNPs, and notably better than the predictions of the ADP. The NNP11 near-neighbor Cu-vacancy and vacancy-vacancy binding energies do deviate from the DFT values by 20 meV, but the ADP predicts massive overbinding of Cu-Cu and Cu-vacancy near-neighbor pairs and underbinding of Cu-Cu third and fourth neighbors. While these ADP values collectively lead to reasonable formation energies for the θ and θ' structures (see Fig. 4), it is evident that such agreement is due to significant cancellations of errors among the different pair interactions. These pair interactions would also have serious negative implications for kinetic studies of precipitation using the ADP.

Table I shows the solute misfit volumes for Cu in Al and Al in Cu, as predicted by DFT, the NNPs, and the ADP. The misfit volumes are computed from the relaxed total volumes of cells containing one solute when using the potentials. When using DFT, we fix the cell size at the bulk elemental value, measure the induced pressure on introduction of the solute, and compute the misfit volume using the pressure and metal bulk modulus. The ADP is grossly incorrect for Cu in Al. This is the origin of reported difficulties in measuring solid-solution strengthening of dilute Al-Cu alloys [26]. The ADP also predicts the incorrect sign for the volume of Al in Cu. The NNPs show a range of values around the DFT value for Cu in Al, with NNP11 in reasonable agreement. The NNPs for Al in Cu are slightly smaller than, but within the range of, the DFT, although NNP11 predicts a particularly small value. Solute strengthening at $T = 0$ K scales as (misfit volume) $^{4/3}$, and hence the deviation for the NNP11 is not negligible for accurate prediction of solute strengthening but other NNPs could yield suitable values for misfit volume but offset by larger errors in the elastic constants.

Table I also shows the solute or stacking fault interaction energies for Cu in Al and Al in Cu, as predicted by DFT, the NNPs, and the ADP. This interaction energy is computed from various system energies as

$$\Delta E_{f,(N-1,X)}^{\text{Sol-SF}} = E_{N-1,X}^{\text{Sol-SF}} - E_N^{\text{SF}} - (E_{N-1,X}^{\text{1sol}} - E_N^{\text{Pure}}), \quad (12)$$

where $E_{N-1,1}^{\text{Sol-SF}}$ is the energy of a system containing the stacking fault and a single solute and E_N^{SF} is the energy of the system with the stacking fault alone. For Cu in Al, the NNPs are in good agreement with DFT while the ADP value is far too large. In contrast, for Al in Cu, the ADP value is in reasonable

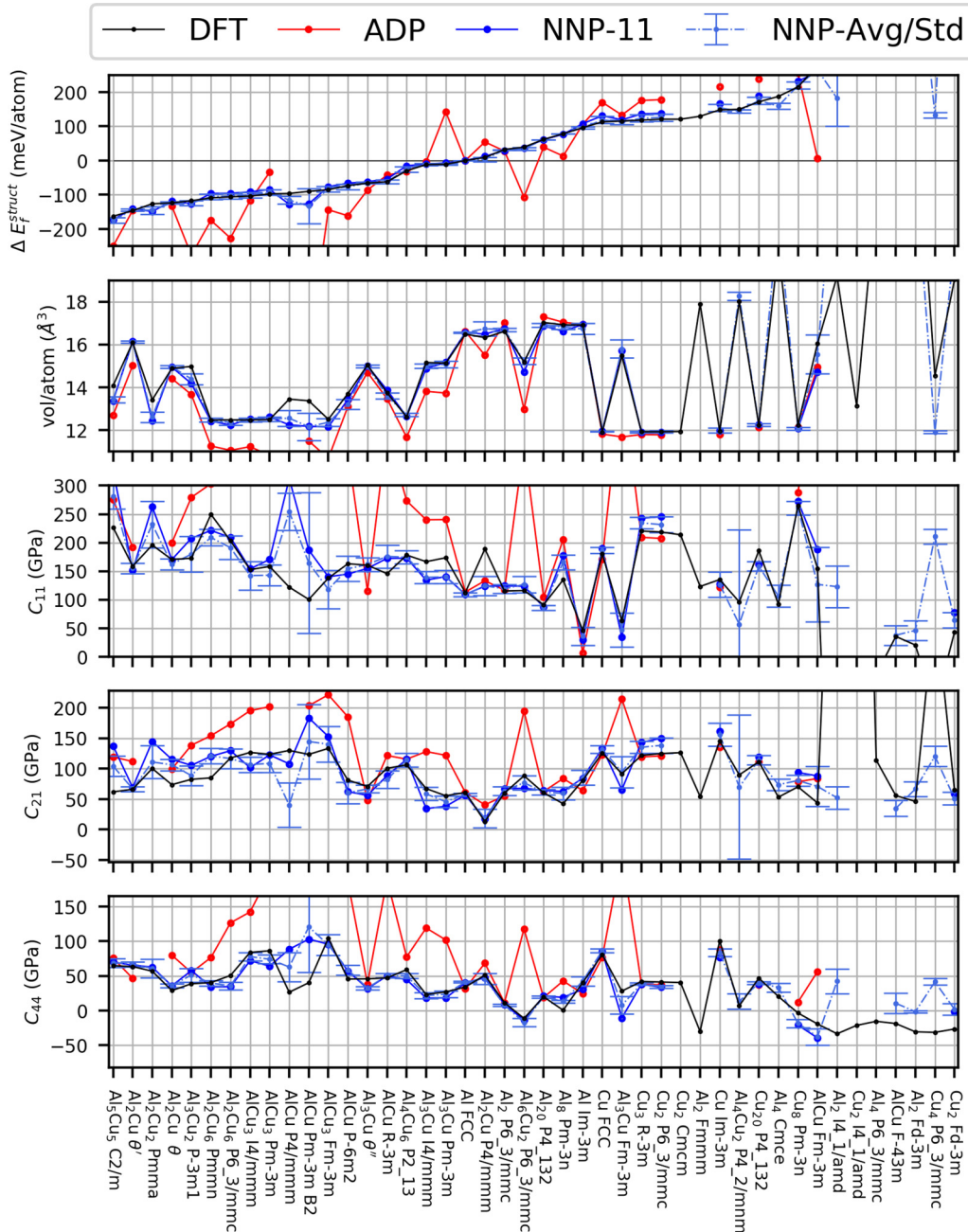


FIG. 4. Formation energy, atomic volume, and elastic constants C_{11} , C_{21} , and C_{44} as predicted by DFT, the NNPs, and the ADP versus structure for all Al-Cu structures in the OQMD database. All structures from which these properties are derived are included in the training set.

agreement with DFT while the NNP values are too small and around zero. This energy can enter into solute strengthening [50] and is crucial for dynamic strain aging (DSA) [51]. With Cu-15at%Al a classic alloy for DSA, the ADP appears better than the NNP but actually remains unsuitable due to the error in the misfit volume.

D. Interface and generalized stacking fault energies

The interface energy between a precipitate and a matrix is an important component of the thermodynamics of nucleation and growth of nanoscale precipitates. In addition, the interaction of a matrix dislocation with a precipitate is

first dominated by the dislocation-interface interactions. It is thus important that any potential provide an accurate description of the precipitate-matrix interfaces. This is challenging due to the absence of information about the interface structure, but machine learning of the energies of approximate structures is valuable for providing some guidance to potentials for modeling of fully relaxed interfaces. Interface energies are almost never considered in the fitting of EAM-type potentials, leading to significant quantitative issues when interface-dominated phenomena are studied.

The interface energy is computed as in Ref. [52] and is briefly summarized here. The formation energy for an interface with N atoms of which X are of the matrix and Y are of

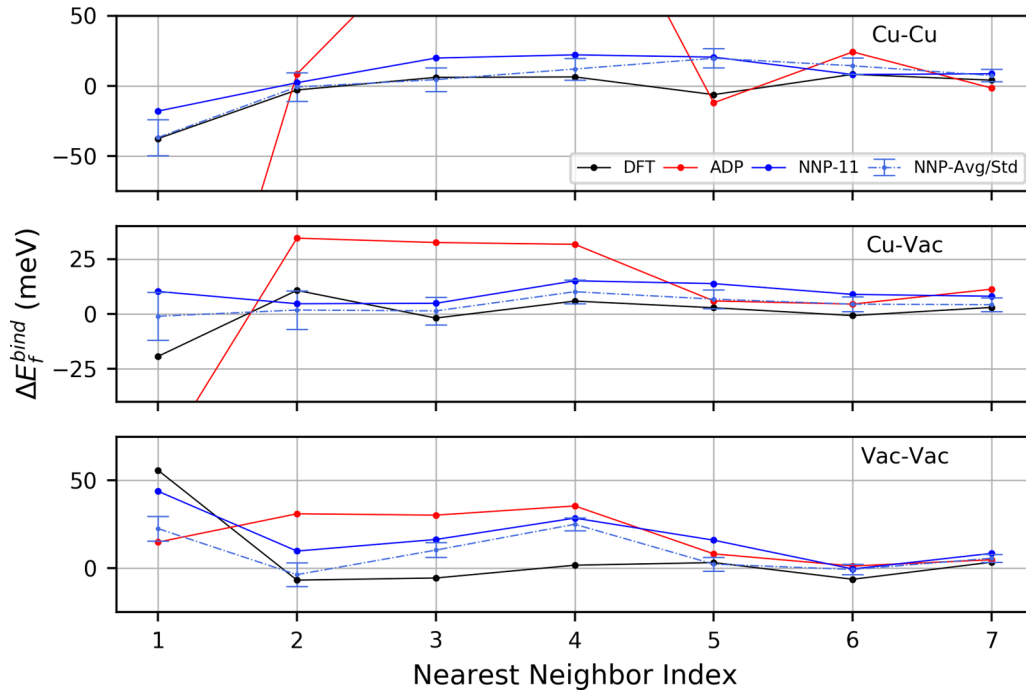


FIG. 5. Pair binding energies for Cu-Cu, Cu-vac, and vac-vac pairs in an Al matrix as a function of near neighbor index. ADP predicts huge errors for Cu-Cu first-neighbor binding (−318 meV) while third and fourth neighbors are far too repulsive (108 and 175 meV, respectively). Relevant structures are included in the training set.

the precipitate, $\Delta E_{f,(X,Y)}^{\text{interface}} = E_{X,Y}^{\text{interface}} - X E_{\text{bulk}}^{\text{Matrix}} - Y E_{\text{bulk}}^{\text{Precip}}$, is related to the number of atoms as

$$\Delta E_{f,(X,Y)}^{\text{interface}} = \delta E_{X,Y}^{\text{strain}} + \frac{2A\gamma^{\text{interface}}}{N}, \quad (13)$$

where $\Delta E_f^{\text{interface}}$ is the interface formation energy, δE_{strain} is the strain energy, A the surface area, γ^{int} the interface energy, and $N = X + Y$ the total number of atoms in interface. During relaxation, atoms are free to move, but the overall cell size and shape are held fixed. The interface surface energy $\gamma^{\text{interface}}$ is then the slope of $\Delta E_f^{\text{interface}}$ versus $\frac{1}{N}$ and divided by $2A$.

Figure 6 shows the computed interface energies from DFT, the NNPs, and the ADP. The ADP correctly predicts the general trends but makes several serious errors. The coherent θ' interface energy is far too high while all the θ'' interfaces are predicted to have negative energy. The θ'' interface is particularly challenging because it has such low interface energy; therefore, it is easy for a potential to predict a negative value. In contrast, NNP11 shows good predictions versus DFT for all of these interfaces. More details on the underlying data for these energies are given in the Supplemental Materials [45].

The GSF energy surface (GSFE) determines the stable faults for the dissociation of dislocations, the core structure of

TABLE I. Misfit volumes and solute stacking fault interactions for DFT, ADP, and NNP.

	DFT	ADP	NNP11	NNP-Avg/SD
Misfit Vol Cu in Al (\AA^3)	-6.077	-16.958	-4.380	-7.095/1.076
Misfit Vol Al in Cu (\AA^3)	2.187	-1.259	0.109	1.334/0.857
SolSF Cu in Al (eV)	0.052	0.108	0.042	0.062/0.030
SolSF Al in Cu (eV)	-0.049	-0.034	-0.006	0.002/0.031

such dissociated partial dislocations, and the energies related to shearing of the material. For precipitates in Al, the GSFE for surfaces best-aligned with the (111) slip planes in Al determine the resistance of the precipitate to shearing and hence control the strength of the alloy. It is thus important that the GSFE well represent all phases that would be used to study dislocation motion and strengthening.

The GSFE is computed using the methods detailed in Ref. [53], which we briefly summarize here. The smallest periodic cell vectors a_1 and a_2 for the desired slip surface are first identified and a supercell created with a_3 normal to the a_1 - a_2 plane. The energy of the undeformed cell E_{pristine} was computed. For each vector \vec{l} of the GSF, the periodic cell

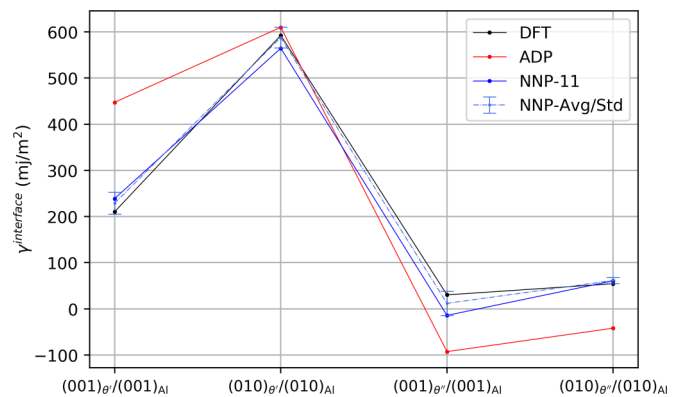


FIG. 6. Interface structure vs. interface energy for precipitate structures. These are the θ' coherent, θ' semicoherent, θ'' coherent, and θ'' semicoherent interfaces in order. Relevant structures are included in the training set. Figure S3 in the Supplemental Material shows the atomic structure of these interfaces [45].

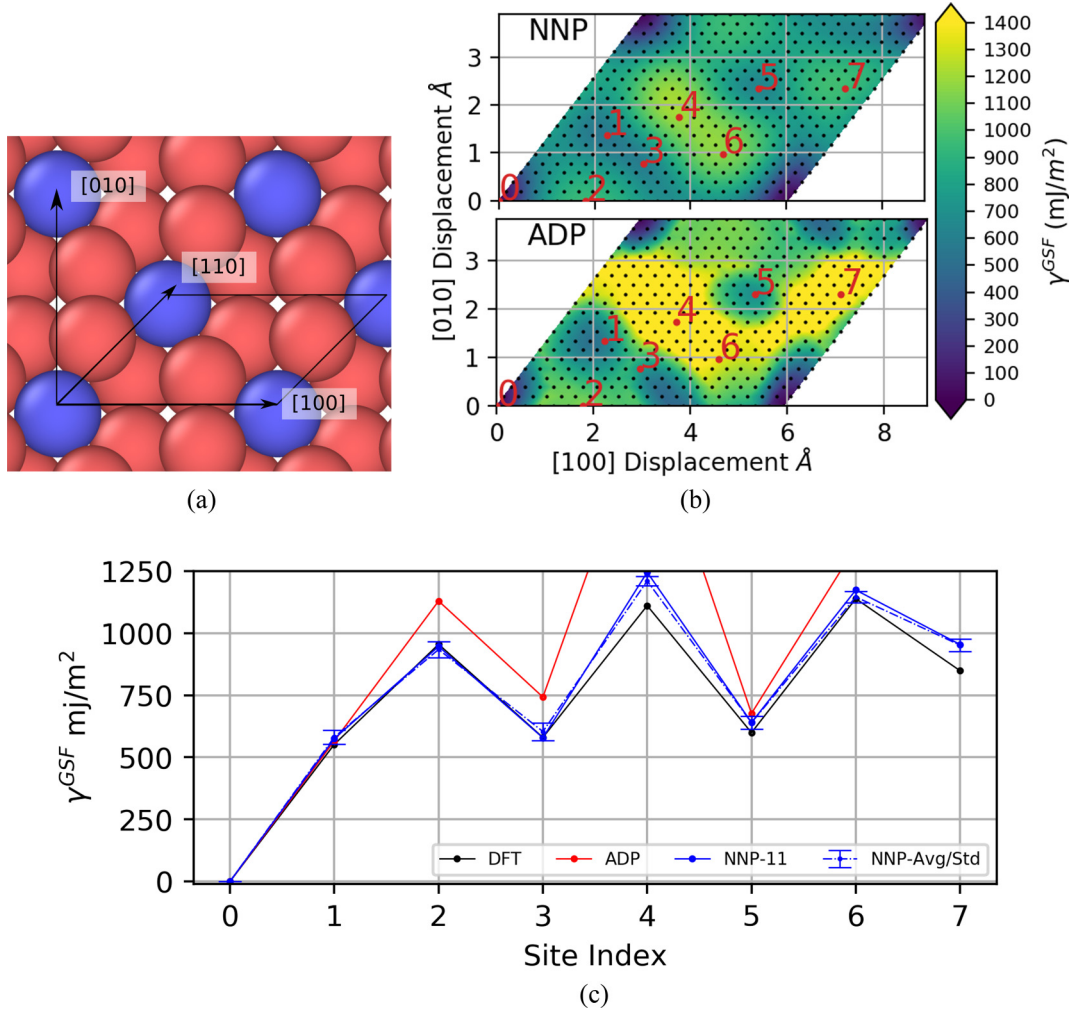


FIG. 7. (a) In-plane unit cell for the $(0\bar{1}1)$ generalized stacking fault surface of the θ precipitate; (b) generalized stacking fault energy surface as computed by NNP11 and the ADP at a discrete set of points indicated by black dots, from which the energy contour is constructed. Red points and labels are local minima or maxima for which the energies were computed via DFT and present in the training set; (c) GSFE at the selected extremal points, as computed by DFT (black dots), all the NNPs (blue dashed lines), NNP11 (blue dots), and the ADP (red dots); the structures contributing to the GSFE for these seven sites are included in the training set.

shape was deformed by shifting a_3 by \vec{r} . All atoms and the a_3 cell vector were then allowed to relax in the 3-direction only to compute the energy $E_{\vec{r}}^{\text{GSF}}$. The GSF energy at a displacement $\gamma_{\text{GSF}\vec{r}}$ was then calculated as

$$\gamma_{\vec{r}}^{\text{GSF}} = (E_{\vec{r}}^{\text{GSF}} - E^{\text{pristine}})/A, \quad (14)$$

where A is the area of the slip surface in the simulation cell.

Figure 7 shows the geometry and GSFE surface for the $(0\bar{1}1)$ for the θ phase as computed by NNP11 and the ADP and the energies at the local minima and maxima as computed by DFT, all the NNPs, and the ADP. The NNPs all predict the extremal values in very good agreement with DFT. In contrast, at the extremal points, the ADP deviates significantly from the DFT for the unstable points 4, 6, and 7 and remains quantitatively much worse than the NNP for the sites 2, 3, and 5. The remainder of the GSFE surface is outside of the DFT training set. The GSFE for the NNP11 is smooth while the ADP shows very sharp gradients that are not expected to be realistic. The normal concern about ML potentials is

that they may have sharp gradients and highly inaccurate energies in regions away from the training set, but the present results demonstrate that it is the traditional ADP potential that suffers from such issues. The GSFE results for the NNP11 indicate that the NNPs can be used reliably for the study of θ precipitate shearing, whereas the ADP cannot.

IV. RESULTS: TRANSFERABILITY

A. Generalized stacking faults θ''

GP zones and θ'' precipitates are among the most pertinent to understanding mechanical properties of Al-Cu alloys [22]. The most important plane of shearing in the precipitate is then that which most closely corresponds to the (111)-type glide planes of Al. Both the θ'' precipitates and the GP zones are structurally similar, being composed of (100) planes of Cu in Al. The θ'' consists of (100) planes of Cu with one intervening plane of Al spaced by two lattice constants while a GP zone is a single plane of (100) Cu. Here we examine the (111) GSFE surface of θ'' since a GP zone does not have a periodic

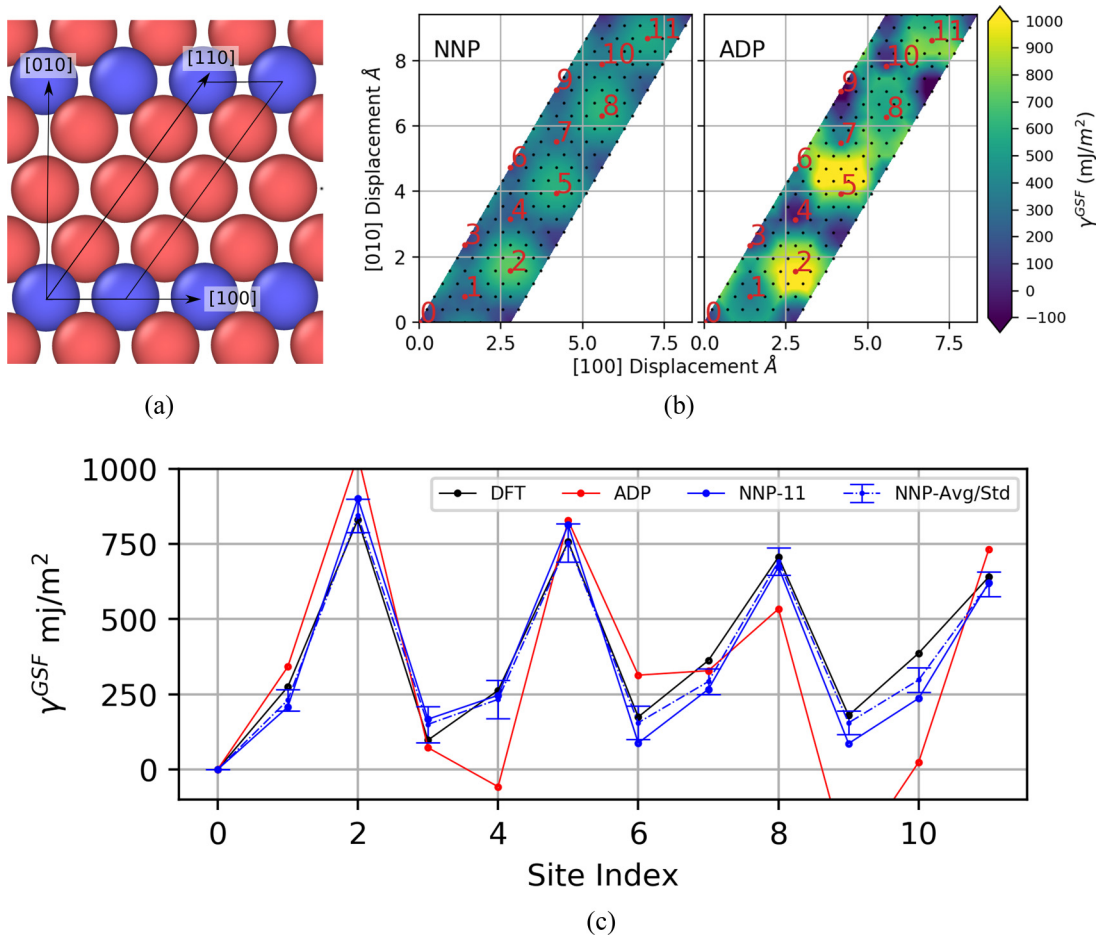


FIG. 8. In-plane unit cell for the (111) generalized stacking fault surface of the θ'' precipitate; (b) generalized stacking fault energy surface as computed by NNP11 and the ADP at a discrete set of points indicated by black dots, from which the energy contour is constructed. Red points and labels are local minima or maxima for which the energies were computed via DFT and present in the training set; (c) GSFE at the selected extremal points, as computed by DFT (black dots), all the NNPs (blue dashed lines), NNP11 (blue dots), and the ADP (red dots); the structures contributing to the GSFE for these 11 sites are not included in the training set.

structure; however, results for θ'' should extrapolate well to GP zones.

Figures 8(a) and 8(b) show the GSFE surface geometry and the energies over the slip surface θ'' , as predicted by the NNP and the ADP, even though outside the training set the NNP predicts a smooth GSFE with maxima and minima at the expected symmetry points. The ADP prediction is much less smooth, with a lack of radial symmetry around the extremal points and with sharp ridges extending across the high-energy regions. Most troubling is that regions of the ADP-predicted θ'' surface are negative, i.e., the sheared precipitate is more stable than the unsheared precipitate. The ADP cannot be used reliably to investigate shearing of these precipitates. Figure 8(c) shows the GSFE energies at the extremal points indicated in Fig. 8(b) as predicted by DFT, the NNPs, and the ADP. The NNPs are in excellent agreement with the DFT results at all points while the ADP is in serious error at several points. Relative to the GSFE of θ (011) (Fig. 7), which was in the training dataset, the accuracy of the NNP is lower and with notably higher variation across the family of NNPs. Nonetheless, the accuracy remains quite good, validating that these NNPs can be reliably used in complex mechanically important environments. We also note

that the NNPs can be further improved by the inclusion of more training data relevant to particular problems of interest, whereas traditional potentials such as the ADP remain high constrained by the limited degrees of freedom available for describing alloys.

B. Antisites and vacancies

In modeling the evolution of precipitate microstructures, it is unavoidable that there will be antisite defects and vacancies inside the precipitates. These configurations also serve as further tests of the ability of the NNPs to broadly capture the energetics of the entire Al-Cu system. For this reason, we compute the unrelaxed antisite and vacancy formation energies for every OQMD structure predicted to be stable using the NNPs (see Fig. 4). Unique atomic sites in each structure were identified using pymatgen. The primitive unit cell was then used to generate larger cells with a minimum of 108 atoms to minimize defect-defect interactions. For each of the unique sites identified, a different atom was substituted, e.g., Al \rightarrow Cu or Cu \rightarrow Al, or the atom was removed to create a vacancy. The unrelaxed formation energies $\Delta E_f^{\text{antisite}}$ of all the antisite structures were then calculated relative to the solid solution

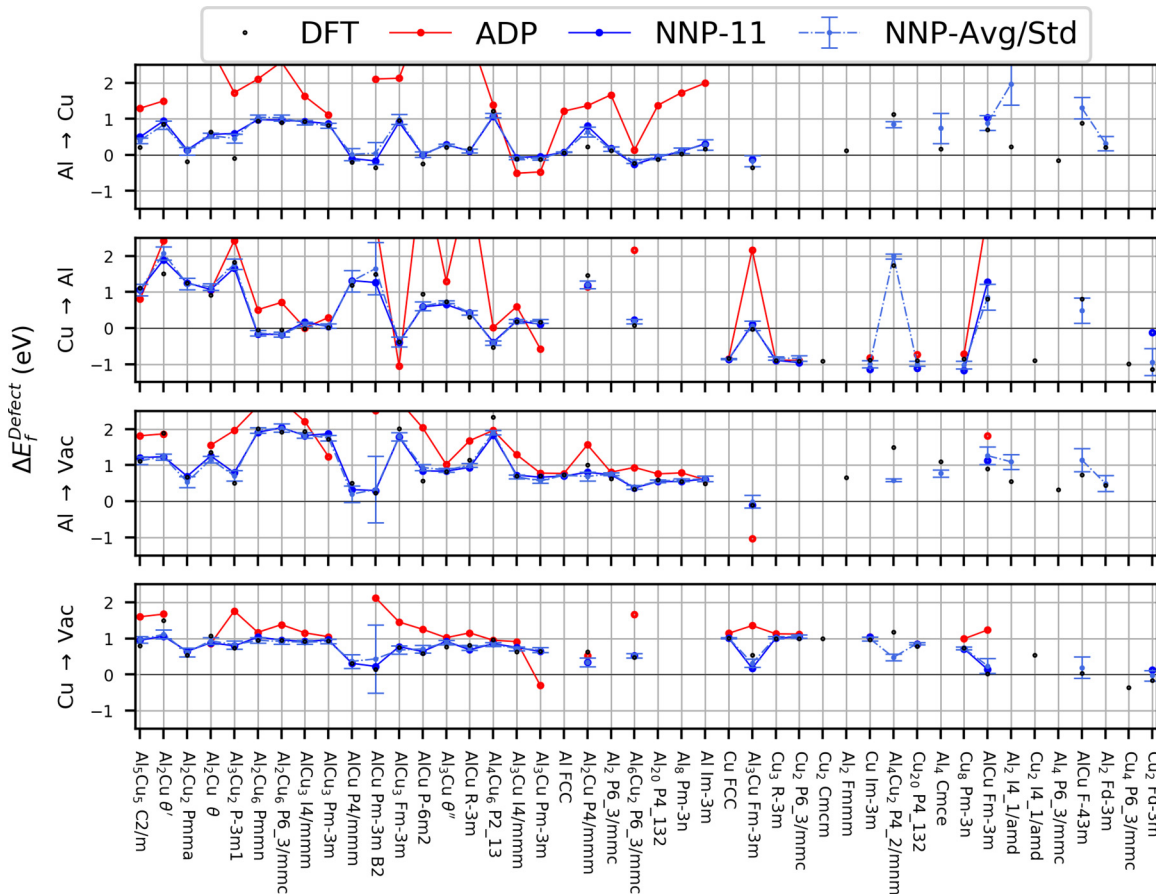


FIG. 9. DFT antisite formation energy vs. ADP or NNP antisite formation energy. These structures were not included in the training set.

alloy as

$$\Delta E_f^{\text{antisite}} = E^{\text{antisite}} - E^{\text{pristine}} - E_{\text{new}}^{\text{ref}} + E_{\text{old}}^{\text{ref}}, \quad (15)$$

where E^{antisite} and E^{pristine} are the total energies of the precipitates with and without the antisite defect and $E_{\text{new}}^{\text{ref}}$ and $E_{\text{old}}^{\text{ref}}$ are the reference energies for the new and old element, respectively, as stated in Eqs. (9) and (10).

Figure 9 shows the antisite results. When there are multiple possible antisites or vacancies for a given element in a structure, we only plot the one with the lowest DFT formation energy; this is purely to make the figure easier to read, the overall accuracy being the same for all vacancies and antisites. For completeness we include Fig. S4 in the Supplemental Material [45], which shows the energies for all antisites and vacancies. Error bars rarely exceed 0.1 eV and are almost always centered around the correct DFT value. The NNP-calculated energies for relaxed sites for Al, Cu, θ , θ' , and θ'' , deviate from their unrelaxed configurations by 0.1 eV, i.e., within the margin of error of DFT vs. NNP. ADP also performs reasonably but has substantially more errors, with many cases deviating by more than 1 eV from DFT. These results demonstrate that NNPs make accurate predictions of point defects that are not in the training set.

C. θ -to- θ' transition temperature

We now turn to the investigation of the phase transition between the θ phase and the θ' , comparing the NNPs to the ADP.

The $T = 0$ K formation energies of both phases have been presented in Fig. 4, and both the NNPs and ADP provide good predictions, with the ADP results also previously reported [25]. An analysis of the finite- T Gibbs free energy starts with the underlying equations of state of both phases. We therefore fit the $T = 0$ K energies to the Vinet equation of state and obtain the zero-temperature bulk moduli K and derivatives of the bulk moduli with respect to pressure K' for θ and θ' . The K values agrees well with the ones derived independently from the elastic constants. As shown in Table II, the ADP potential predicts a large negative value for K' of the θ phase, which has critical implications for phase stability (see below). The K' values predicted by NNP11 differ from the DFT references but are positive for both phases and in qualitative agreement with DFT.

To obtain the full temperature dependence of the Gibbs free energy including all anharmonic contributions for both the θ' and θ phases, we follow the strategy developed in

TABLE II. Derivative of the isothermal bulk modulus K' at $T = 0$ K for the θ - θ' phases comparing ADP, DFT, and our NNP11 potential.

Phase	ADP (this work)	DFT	Neural network
Theta (θ)	-4.2	4.55	2.29
Theta prime (θ')	2.9	4.58	1.03

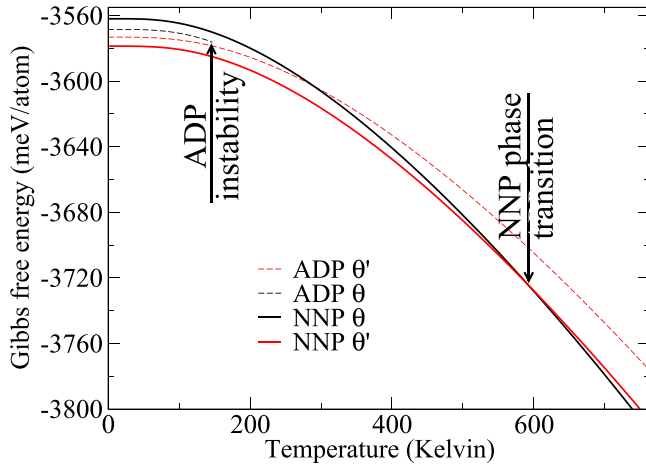


FIG. 10. Gibbs free energy versus temperature for the θ and θ' phases of Al-Cu as predicted using NNP11 and the ADP. The ADP predicts an instability for the θ phase at ≈ 150 K while NNP11 predicts no instabilities and a phase transition at just below 600 K. Note that for plotting purposes, a constant energy has been added to the NNP11 Gibbs free energies.

Refs. [54–56]. We first compute the harmonic vibrational Helmholtz free energy,

$$F^{\text{vib}}(V, T) = \frac{1}{N} \sum_i^{3N} \left[\frac{\hbar\omega_i}{2} + k_B T \ln[1 - e^{-\beta\hbar\omega_i}] \right], \quad (16)$$

at seven atomic volumes V over the range needed for the temperatures considered. Here N is the number of atoms, k_B Boltzmann constant, and $\beta = (k_B T)^{-1}$. We investigate supercells of up to $N = 768$ (for θ) and $N = 648$ (for θ') and obtain the $3N$ eigenvectors \mathbf{k}_i and eigenfrequencies ω_i by direct diagonalization of the force constant matrix. The quasiharmonic free energy surface $F(V, T) = E(V) + F^{\text{vib}}(V, T)$ is minimized with respect to the volume to define the zero-pressure volume at each temperature and obtain the corresponding zero-pressure quasiharmonic Gibbs free energy. Finally, we perform thermodynamic integration [55,57] from the (quasi)harmonic reference to the *full* Gibbs free energy surface including anharmonic contributions. Convergence tests show that supercells containing 96 and 81 atoms are sufficient to reach an accuracy in the phase transition temperature to within 10 K.

Figure 10 shows the Gibbs free energy versus temperature of both phases, as predicted for both NNP11 and the ADP. Notably, the ADP θ phase becomes unstable above 150 K so that no transition from θ' to θ can be obtained. A previous study of the θ phase using the ADP in a quasiharmonic analysis with linear expansion of the Gibbs free energy versus volume [25] suggested a stable θ phase, which is now seen to be incorrect. In contrast, NNP11 shows a smooth variation in the Gibbs free energies for both phases, with the phase transition predicted to be just below 600 K. Across the 40 different NNPs, we obtain a transition temperature of 596 K with a standard deviation of 5 K, showing high consistency across all NNPs. As an aside, the average quasiharmonic transition temperature is 31 K lower at 565 K. Very early experimental data for this transition reported a transition temperature of 463 K [58].

The fact that the NNPs predict the transition at a temperature in reasonable agreement with experiment is very positive, especially recalling that the NNPs were only trained on the $T = 0$ structures and various Al-Cu structures with random imposed displacement. An error in the relative free energy of just a few meV would shift the predicted transition temperature by 100 K, and the overall training of the NNPs is only accurate to within this energy range. This analysis thus provides further confidence in the transferability of the NNPs developed here.

D. Solid-to-liquid transition temperature

We have investigated the solid-to-liquid transition temperature of NNP11 for pure Al and for a near-eutectic Al-18%Cu solid solution using molecular dynamics simulations using LAMMPS. We follow the solid-liquid coexistence methodology as follows. One half of the atoms in a periodic simulation cell are first equilibrated at a temperature well above the melting point and the other half equilibrated below the melting point, while maintaining zero pressure on the entire cell. The equilibration time is 10 ps with a time step of one fs. The entire system is then evolved further within the Isoenthalpicisobaric ensemble (NPH) ensemble at zero pressure for another 50 ps. The solid-liquid coexistence temperature (melting point) is found as the temperature at which the interface position is stable during the latter portion of the simulation.

Figure 11 illustrates the solid-liquid coexistence geometries at different average system temperatures for pure Al. At $T = 800$ K the liquid region is shrinking and the solid growing, while at $T = 850$ K the solid region is shrinking. The interface was stable at 825 K for 10 ps, with a slight preference for the liquid region. Therefore, the estimated melting temperature of pure aluminum for NNP11 is $T_m = 825$ K approximately 108 K lower than the experimental melting point (933 K). A similar study using potential NNP10 gave a similar result of T_m 840 K.

Figure 12 shows results for a similar study at the near-eutectic composition Al-18%Cu. The eutectic point is an equilibrium between the liquid and two crystalline phases (a dilute solid solution and an intermetallic). We examine the equilibrium between the liquid phase and the crystalline solid solution at the same composition; this temperature should be below the eutectic point. Simulations with NNP11 show that the liquid phase is stable down to 600 K, with a transition to the crystalline solid-solution structure at an estimated temperature of $T = 560$ K. This transition temperature is lower than the experimental eutectic point of 823 K, as expected. The main feature here is thus that the Al-18%Cu liquid is stable at much lower temperatures than pure Al, as seen experimentally. A similar study repeated using potential NNP10 gave a similar result of 600 K.

E. Solute clusters and kinetic Monte Carlo

In early-stage precipitation, preprecipitates are nucleating by formation of small multiatom clusters. An understanding of the nucleation and early-stage growth thus requires that a potential properly represent not only many possible precipitate phases and precipitate-matrix interfaces but also

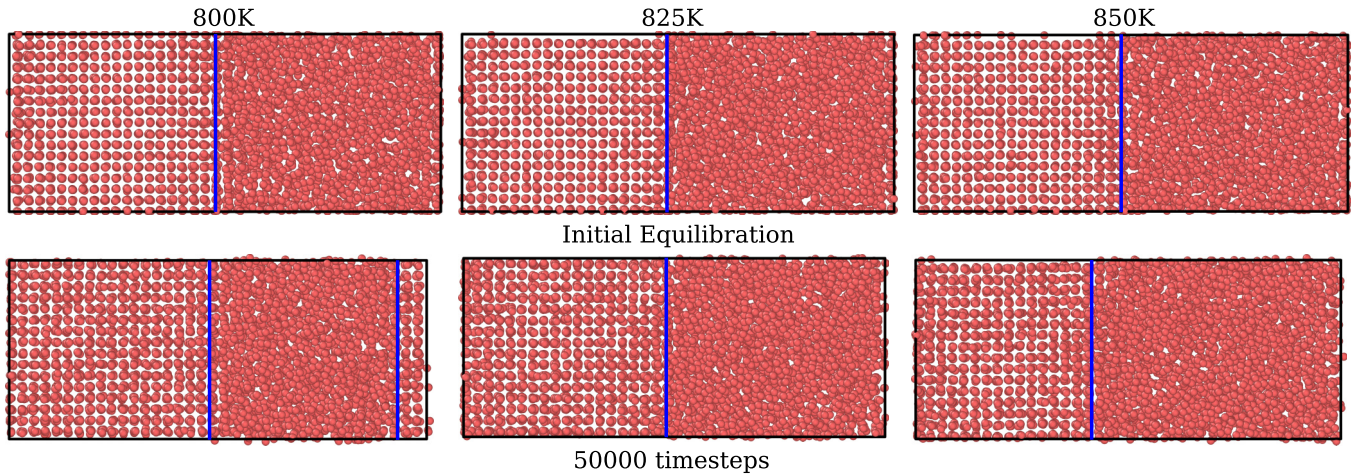


FIG. 11. Solid-liquid coexistence in pure Al at different temperatures simulated using the NNP11 potential. The blue lines indicate the solid-liquid interface.

possible clusters that cannot be classified as precipitates. The pair interactions have already been discussed, and so here we focus on several three- and four-atom clusters of Cu in Al that have recently been studied by Gorbатов *et al.* [59] in their development of a cluster expansion method. The specific clusters are illustrated in Table III and we use the notation of Gorbатов *et al.* Among these clusters, we note that the 111 and 111111 (compact triangle and tetrahedron of nearest-neighbor Cu atoms, respectively) are not compatible with the formation of GP-I and GP-II zones on the (100) planes that occur in real Al-Cu while the 112, 111112, and 111122 clusters are consistent with GP-I and GP-II zones.

We compute the cluster formation energy $\Delta E_{f,(N-X,X)}^{\text{cluster}}$ as the energy of the embedded cluster relative to the energy of Cu in solid solution,

$$\Delta E_{f,(N-X,X)}^{\text{cluster}} = E_{N-X,X}^{\text{cluster}} - \sum_i E_i^{\text{ref}}. \quad (17)$$

The DFT values for all the clusters as computed in simulation cells of 256 atoms are shown in Table III along with the NNP and ADP values. The DFT predicts that clusters favoring the

(100) plane are energetically preferable as compared to the other clusters and with the 111111 tetrahedron cluster having positive formation energy. NNP11 is in reasonable agreement with DFT for most of the clusters but notably incorrectly predicts a strongly negative formation energy for the 111111 cluster and also overpredicts the stability of the 111 cluster. Shown for comparison are the results from NNP10 which is in better agreement with DFT with respect to the relative stability of the clusters. Also shown are the results from NNP13, which performs better for small three- and four-atom clusters but deviates at larger sizes. Across the mean and standard deviation of the 40 NNPs, the deviations seen for NNP11 are also not evident. The ADP potential shows gross errors, predicting all clusters to be very energetically favorable and predicting the relative stability of the three-atom clusters incorrectly (as does NNP11 but with a much smaller difference). Application of ADP to the study of early-stage clustering is thus expected to be very inaccurate while the performance of the NNPs may vary. We will investigate this next.

To examine the early-stage clustering under realistic conditions of time and temperature, we perform off-lattice KMC

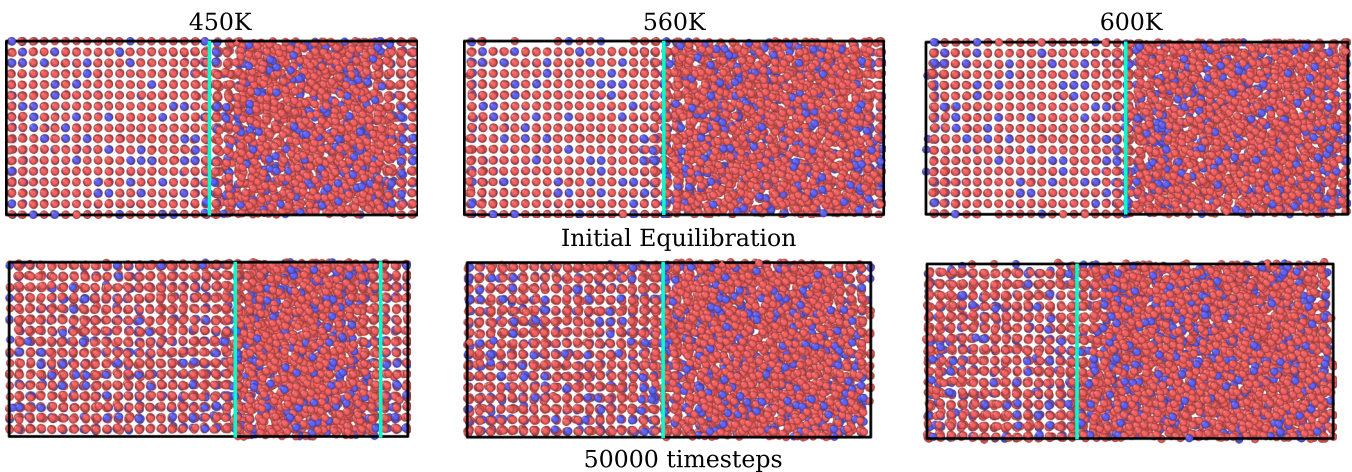









FIG. 12. Solid-liquid solution coexistence in Al-18%Cu at different temperatures simulated using the NNP11 potential. The cyan lines indicate the solid-liquid interface.

TABLE III. Cu cluster energies for DFT, ADP, and NNP. Clusters are described using the length of edges of the cluster, e.g., 111 represents three edges that are all nearest neighbors. Note that there are three edges for a three-atom cluster but six edges for a four-atom cluster.

	DFT	ADP	NNP11	NNP10	NNP13	NNP-Avg/ <i>SD</i>	Figure
$E_{111}^{3\text{Cu}}$ (eV)	-0.073	-0.718	-0.112	-0.096	-0.081	-0.060/0.037	
$E_{112}^{3\text{Cu}}$ (eV)	-0.112	-0.595	-0.097	-0.127	-0.110	-0.085/0.025	
$E_{11111}^{4\text{Cu}}$ (eV)	0.039	-0.612	-0.081	-0.003	-0.016	0.071/0.086	
$E_{11112}^{4\text{Cu}}$ (eV)	-0.126	-1.071	-0.184	-0.183	-0.185	-0.093/0.066	
$E_{11122}^{4\text{Cu}}$ (eV)	-0.268	-1.275	-0.226	-0.312	-0.303	-0.205/0.055	
$E_{\text{Unrelaxed}}^{10\text{Cu}}$ (eV)	1.62	5.141	1.589	2.254	0.967	2.275/0.618	
$E_{\text{relaxed}}^{10\text{Cu}}$ (eV)	-	-0.606	-0.043	0.961	-0.512	0.700/0.589	

simulations of the evolution of an Al-(4-5)at% Cu alloy starting from a random solid solution state. For this demonstration, we use the bulk vacancy migration barrier $E_{\text{Al}}^{\text{Mig}}$ in Al (computed from DFT) as the migration barrier for all vacancy-mediated transitions and compute the enthalpy barrier between the initial state i and final state f using a simple chemical-kinetics model as

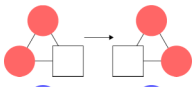
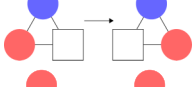
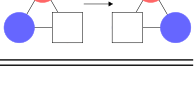
$$E_{i-f}^{\text{Mig}} = E_{\text{Al}}^{\text{Mig}} + (E_f - E_i)/2, \quad (18)$$

where E_f and E_i are the fully relaxed final and initial state energies of the entire system. Table IV shows the migration barriers for the three main transitions in a dilute alloy: the Al-V exchange in bulk, the Al-V exchange in a triplet with Cu neighbor, and the Cu-V exchange, as computed via DFT [60], ADP, NNP11, and NNP10, which will be discussed

further below. All three migration barriers are well captured by both NNPs but the approximation above will use only the bulk vacancy migration barrier for all transitions, including those in and around small clusters the form during the KMC. Thus, the KMC will satisfy detailed balance and capture the correct thermodynamics but not represent the true timescale of evolution [60] of the system. We use a 1000-atom supercell containing one vacancy and execute simulations at 300 K with 4at% Cu and at 696 K with 5at% Cu using both NNP11 and NNP10, the latter providing a better representation of the energetics of the three- and four-atom Cu clusters shown in Table III.

Figure 13 shows KMC simulations using NNP11 and NNP10 with Al-4at% Cu at 300 K after 5×10^5 KMC steps. Both systems ultimately form Cu GP-I precipitates consistent

TABLE IV. Vacancy migration barriers with a single Cu atom in the Al matrix.

Jump type	Migration energy (eV)						
	DFT ^a	ADP	NNP11	NNP10	NNP13	NNP-Avg/ <i>SD</i>	
Bulk Al-V	0.58	0.63	0.58	0.58	0.60	0.58/0.01	
Triplet Al-V	0.38	0.46	0.41	0.42	0.43	0.42/0.02	
Cu-V	0.57	0.76	0.43	0.52	0.42	0.47/0.04	

^aThe reported DFT results were computed by Mantina *et al.* [60].

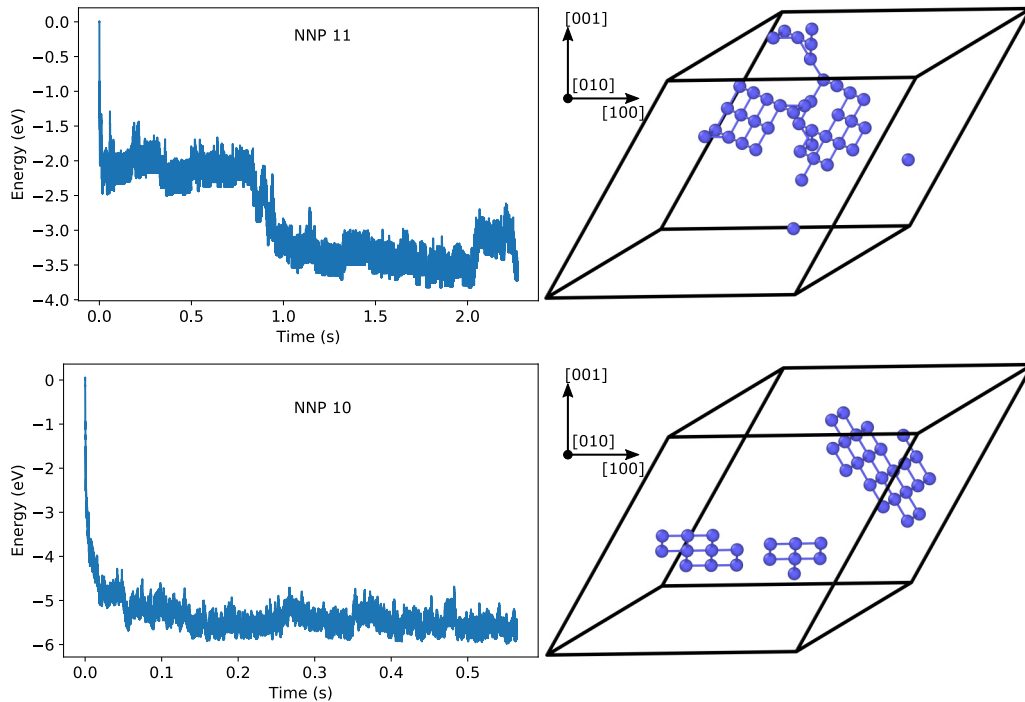


FIG. 13. Energy vs. simulation time and the final geometry showing Cu GP-I precipitate formation. The simulations were carried out with NNP10 and NNP11 at 300 K using a 1000-atom supercell containing 40 Cu atoms for 5×10^6 KMC steps.

with real alloys under similar conditions. NNP11 requires a much longer time, nearly 1 s, before GP-I formation as compared to NNP10 (less than 0.1 s), which may reflect the different competing energetics of the various three- and four-atom clusters in NNP11 as compared to NNP10. The NNP10 structure also has a much lower energy, indicating a more-well-formed precipitation state, relative to NNP11, again indicating that the GP-I formation with NNP11 may compete energetically with some other structures at the early stages of formation. Starting the KMC simulation from different initial random geometries and/or other NNPs leads to the formation of GP-I within similar times and energy levels. The NNPs thus consistently capture the behavior expected in real materials. Gorbatov *et al.* [59] and Miyoshi *et al.* [61] have recently presented KMC studies of the same problem, showing similar formation of the GP-I zones, and so our results are not unique. However, these potentials used cluster expansion methods that, while suitable for KMC, cannot be used for any other studies beyond Al-Cu thermodynamics such as plasticity and fracture studies. The NNPs show the expected precipitation behavior while also being valid for a wide range of applications.

KMC simulations at higher temperature (696 K) [and slightly higher Cu concentration (5at%)] reveal pitfalls of some of the NNPs, however, as can be seen in Fig 14. For the case of NNP13, a dense block of FCC Cu is formed with a very low energy 14 compared to the initial solid solution. This spurious behavior arises because NNP13 has a negative formation energy for the three-atom 111 and four-atom tetrahedral 111111 clusters (Table III) and even larger formation energies for larger dense clusters (10 atom cluster shown in Table III). The formation energy becomes progressively lower as more Cu atoms cluster, ultimately favoring growth

of the dense FCC Cu. Neither NNP11 nor NNP10 exhibit this erroneous behavior, and the KMC simulations at 696 K show no precipitation and only sporadic few-atom clusters of Cu; this is consistent with expectations for real Al-Cu under these conditions. Note, however, that even NNP13 correctly predicts Cu GP-I precipitation at room temperature.

The above findings reinforce that careful validation of any potential (NNP or otherwise) must be done within the range of structures for which the NNP will be used. This result also highlights the use of multiple NNPs for the same problem, especially extracting new structures from a simulation using NNP-X and evaluating their energies using other NNPs as

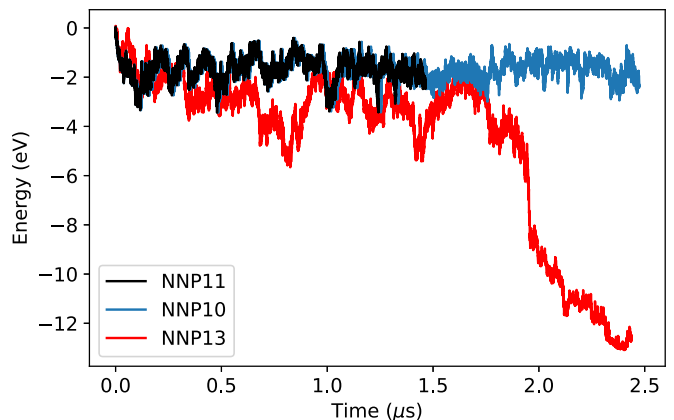


FIG. 14. Energy vs. simulation time for the KMC simulations carried out with NNP10, NNP11, and NNP13 at 696 K using a 1000-atom supercell containing 50 Cu atoms. The NNP11 trajectory was stopped at 3×10^6 steps as it closely follows NNP10, while the NNP10 and NNP13 trajectories were continued for 5×10^6 steps.

some measure of reliability of the NNPs for these new structures. Finally, having a family of NNPs can enable mixing and matching of the NNPs in a set of related studies. For instance, NNP10 can be used for precipitate evolution in KMC while NNP11 (or some other choice) is then subsequently employed to study dislocations moving through the microstructure that was established using NNP10. The close agreement among the family of NNPs for many properties makes such a mix-and-match strategy feasible in a manner that is not possible using traditional potentials.

F. Dislocations and fracture in aluminum

In any applications of potentials to plasticity and fracture, it is essential to demonstrate that the potential reproduces a reasonable structure for the relevant dislocation(s) and exhibits physically expected behavior at sharp crack tips. There are no precise first-principles results for comparisons, but DFT results using other settings and novel simulation methods provide an approximate reference for determining whether the ML potential is reasonable and physical. An ML potential can also be compared to other traditional potentials where the behavior has previously been found reasonable.

With that background, we examined the predictions of the NNP11 potential for the structure and energetics of the $a\langle 110 \rangle/2$ edge and screw dislocations in pure Al. The methodology to create the dislocation is standard and described in Ref. [62]. An edge or screw dislocation is placed at the center of a cylindrical cell of radius 143.1 (Å). All atoms in the cell are deformed according to the anisotropic displacement field of a Volterra dislocation computed using the Stroh formalism [63]. Atoms within 22 Å of the outer boundary are then fixed at these initial positions, and all interior atoms relaxed to a minimum energy condition. The core structure for both edge and screw dislocations for NNP11 is shown in Fig. 15 along with the DFT-computed structure from Ref. [64]. The NNP11 dislocation dissociates into two partial dislocations, as expected. The dissociation distance between the two partials differs slightly from the DFT reference, recalling that the NNP11 has a higher C_{44} than usually obtained in DFT and that this DFT structure is obtained by very different methodology and DFT parameters.

As a quantitative measure, we compute the core energy Γ^{core} energy of the dislocation as defined at distance $r_c =$

$$K_{\text{Ie}} = \begin{cases} \sqrt{\gamma^{\text{USF}} \alpha(\theta^{\text{slip}}, \phi^{\text{Burgers}}) / F_{12}(\theta)} & \gamma^{111} \leq 3.45 \gamma^{\text{USF}} \\ \sqrt{(0.5 \gamma^{\text{USF}} + 0.145 \gamma^{111}) \alpha(\theta^{\text{slip}}, \phi^{\text{Burgers}}) / F_{12}(\theta)} & \gamma^{111} > 3.45 \gamma^{\text{USF}} \end{cases}, \quad (19)$$

where γ^{USF} is the unstable stacking fault energy, γ^{111} is the surface energy of the emission plane, $\alpha(\theta^{\text{slip}}, \phi^{\text{Burgers}})$ is an anisotropic elasticity term dependent on the inclination of the slip plane and Burgers vector, and $F_{12}(\theta^{\text{slip}})$ provides the resolved slip on the slip plane oriented at an angle with respect to the crack plane (see Andric and Curtin [67]). Using the DFT-computed values of the various material quantities, the critical value is predicted to be $K_{\text{Ie}} = 0.341 \text{ MPa} \sqrt{m}$. A multiscale DFT simulation using the same crack orientation but

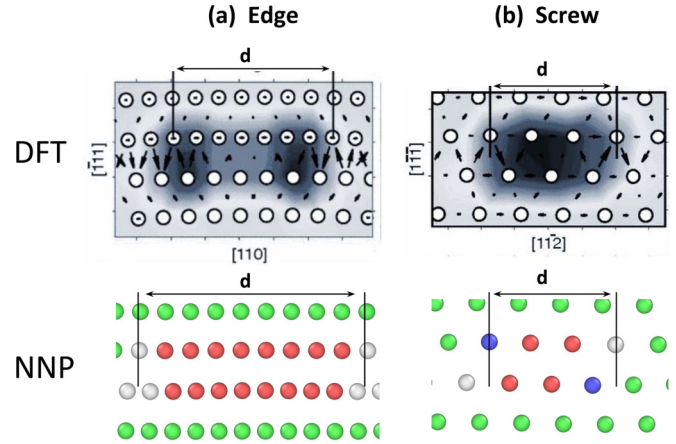


FIG. 15. Atomistic structures of the edge and screw $a\langle 110 \rangle/2$ dislocations in pure Al as predicted by DFT and NNP11. (a) Edge dislocation: DFT shows the differential displacement map and the estimated partial dislocation spacing; NNP shows atoms identified by common neighbor analysis (green, FCC; red, HCP; white, other) where the HCP atoms lie in the stacking fault between the two partials and the partial core regions appear as white, with the estimated partial dislocation spacing shown. (b) Screw dislocation: as in (a) but where white and blue atoms identify the screw partial cores. DFT image adapted from [64] with permission of the APS

using the analysis of Ref. [62]. We compute $E^{\text{core}} = 1.144 \text{ eV/nm}$ for the edge dislocation and 1.055 eV/nm for the screw dislocation, respectively, which compare well to the values from EAM potentials of Ercolessi-Adams [65] (edge: 0.463 eV/nm ; screw: 0.810 eV/nm) and Mishin [66] (edge: 1.198 eV/nm ; screw 1.262 eV/nm) [62]. The EAM potentials have been widely used for modeling of dislocations in Al, and so the values obtained for the NNP, along with the accurate structures, indicate the suitability of the NNP for studying dislocations.

Our broad experience with both ML and traditional potentials has indicated that they can struggle to correctly capture behavior at the tip of a sharp crack. So here we examine the most basic fracture test for an FCC metals, the emission of a dislocation from a sharp crack tip. Theoretical models predict that the critical stress intensity factor K_{Ie} for emission of the first partial dislocation from the tip of a crack with crack plane (111) and crack line direction (110) is

different DFT parameters [68] was shown to emit a dislocation at $K_{\text{Ie}} = 0.368 \text{ (MPa} \sqrt{m})$. We note further that emission occurs rather than cleavage because the stress intensity for cleavage, related to the surface energy, is larger than K_{Ie} . This background provides a basis for assessing the NNPs and demonstrates the range of values that can emerge from DFT assessments.

We simulate the crack tip behavior following the K -test methodology carefully described by Andric and Curtin [67].

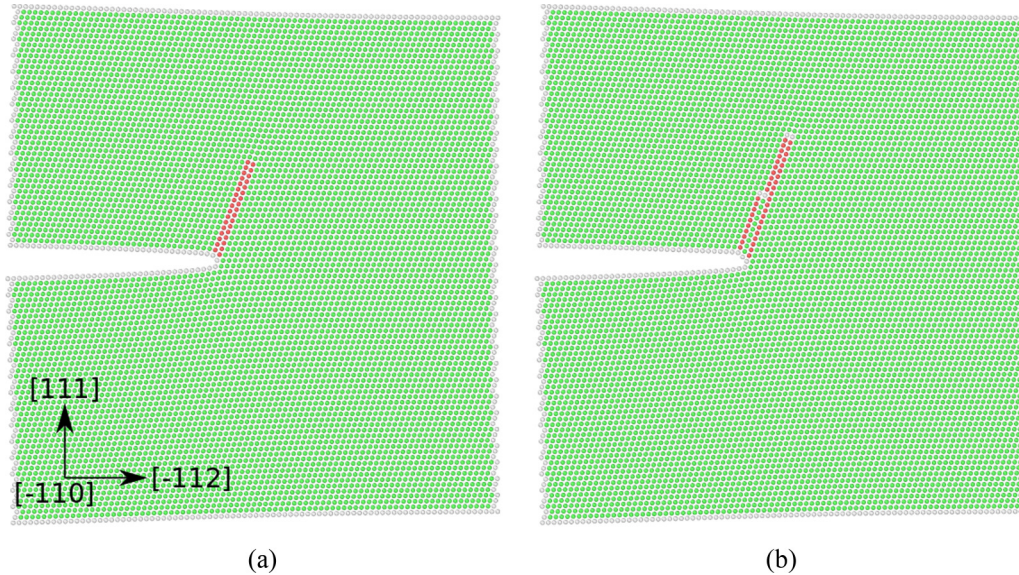


FIG. 16. Observed dislocation emissions: (a) leading and (b) leading and twin partial. Green atoms are those in FCC configuration, red are for HCP (i.e., stacking faults), and white atoms are neither HCP nor FCC (e.g., surfaces, crack tips).

We use a simulation cell of size $200 \times 200 \times 10 \text{ \AA}^3$, insert a crack terminating along the z direction in the center of the cell, hold boundary atoms fixed at the anisotropic elastic predictions for the desired applied stress intensity factor, and allow all interior atoms to relax using the conjugate-gradient method followed by steepest descent with force tolerance $1 \times 10^{-5} \text{ eV/\AA}$. Successive increments in load level are applied until any notable event occurs at the crack tip. All 40 NNPs show partial dislocation emission and with no gross artifacts at the crack tip; a typical example is shown in Figure 16, which is qualitatively similar to results found with many traditional potentials. The emission event occurs at a critical load of $K_{Ie} = 0.354 \pm 0.007 \text{ (MPa}\sqrt{m})$ across all 40 NNPs, in very good agreement with the DFT-predicted theoretical value and consistent with the value found in the direct multiscale DFT simulation. Some variation does arise among the NNPs: Thirteen NNPs showed emission of only the leading partial [Fig. 16(a)] while 27 NNPs showed the emission of a second twinning partial just behind the crack tip [Fig. 16(b)] following the leading partial with no increase in the applied load. The NNP simulations are thus consistent with theory and other simulations showing that this second partial emission event is the next expected event. The load for this second event depends on several additional factors including the unstable twinning fault energy and results from EAM Al potentials show that the load is only slightly higher than that required for the first event. The NNP behavior is thus within the scope of reasonable behavior. Overall, the NNPs for pure Al provide a very good representation of the crack tip and the first partial emission process as compared to DFT, theory, and other simulations.

V. CONCLUSION

We have developed a family of Behler-Parinello-type NNPs for the Al-Cu system based on a comprehensive metallurgically relevant training dataset based on first-principles

DFT. We have demonstrated the ability of these potentials to predict many different properties of importance in metallurgy, including the many possible phases, solute properties, shearing of precipitates, early stage clustering during annealing, dislocation structures, and crack tip performance. We emphasize that it is essential to study the performance of any potential across such a broad spectrum of structures if the potential is to be used for realistic studies of atomistic phenomena. Our examination of the NNPs has included properties that are directly derived from the training dataset and hence expected to be well captured by the non-linear regression of the NN method. But, moreover, we investigated a number of properties and structures that were intentionally not in the training dataset. The methodology is easily extendable to the inclusion of more training data when the current family of NNPs is found to be inaccurate for some problems, as will undoubtedly be the case in the future. To facilitate such future advances, our entire DFT training set is available along with the DFT provenance necessary to expand the database using the same DFT parameters.

The NNPs developed here have been extensively compared to the newest and widely used traditional interatomic potential for Al-Cu, the angularly dependent potential from Apostol and Mishin. While the ADP is based on some underlying physical principles, it has far fewer fitting parameters and hence performs comparably to the NNPs only for those few properties to which the ADP is fit. Across the much broader range of properties that are necessary for realistic metallurgical studies, the NNPs here are demonstrably far better, and the ADP gives unphysical behavior for some important situations. We reiterate that our purpose is not to denigrate the ADP in particular but rather to show that one of the best traditional alloy potentials has many deficiencies and that a machine learning approach overcomes those deficiencies. The overarching philosophical concern about the use of ML potentials is that they contain no physics and so can possibly fail very badly in situations where a traditional potential might also

fail quantitatively but not qualitatively. With our comparison of the NNP to ADP, this concern can be partially relieved. Across a much wider spectrum of properties that must be well represented for realistic atomic-scale studies in alloys, the NNPs perform far better than the ADP and so it is, in fact, use of the ADP that will more likely introduce hidden errors across a wide scope of possible applications, leading to incorrect predictions either qualitatively or quantitatively. A compromise strategy, emerging in the literature, is to combine a traditional potential such as the ADP with an ML potential that is trained on the difference between the traditional potential and a DFT training set. Such a strategy can be pursued in the future for Al-Cu using our openly available DFT training set.

The family of NNPs is not perfect. Most notably, C_{44} of Al has substantial errors even with careful inclusion of relevant structures. Across our family of NNPs, and adding structures intended to bias the optimization toward a better representation of C_{44} , we could not achieve an error below 10%. Similarly, we could not reduce the errors on solute pair energies below 20 meV, again despite the careful inclusion of structures in the training set. These limitations might be resolved through the use of other ML methods, other optimization methods based on energy differences between

structures rather than absolute energies, a larger set of descriptors for the atomic environments, and so on, and all of these can also be pursued in the future using our openly available DFT training set.

In summary, we have demonstrated that a machine learning interatomic potential can be created that accurately captures the exceptionally rich and diverse phase space of a binary alloy, here Al-Cu, with very good and physical performance for material properties and structural response that are necessary for metallurgical studies. This demonstration points to continued development of ML-based interatomic potentials based on an extensive database of carefully chosen metallurgically relevant structures as a viable path toward the quantitative study of metal alloys at the atomic scale.

ACKNOWLEDGMENTS

The authors acknowledge support for this work by the NCCR MARVEL, funded by the Swiss National Science Foundation. The authors acknowledge E. Mak for the fracture results, Y. Hu for the dislocation results, Dr. B. Yin for helpful discussions regarding DFT calculations, R. Ahmad for insight into GSF structures, and Drs. G. Pizzi and S. Huber for their help in employing AiiDA.

-
- [1] N. Juslin, P. Erhart, P. Träskelin, J. Nord, K. O. E. Henriksson, K. Nordlund, E. Salonen, and K. Albe, Analytical interatomic potential for modeling nonequilibrium processes in the W-C-H system, *J. App. Phys.*, **98**, 123520 (2005).
 - [2] A. P. Bartók, R. Kondor, and G. Csányi, On representing chemical environments, *Phys. Rev. B* **87**, 184115 (2013).
 - [3] J. Behler, Constructing high-dimensional neural network potentials: A tutorial review, *Int. J. Quant. Chem.* **115**, 1032 (2015).
 - [4] R. Kobayashi, D. Giofré, T. Junge, M. Ceriotti, and W. A. Curtin, Neural network potential for Al-Mg-Si alloys, *Phys. Rev. Mater.* **1**, 053604 (2017).
 - [5] G. P. Purja Pun, R. Batra, R. Ramprasad, and Y. Mishin, Physically-informed artificial neural networks for atomistic modeling of materials, *Nat. Commun.* **10**, 2339 (2019).
 - [6] Q. Wang, Z. Li, S. Pang, X. Li, C. Dong, P. Liaw, Q. Wang, Z. Li, S. Pang, X. Li, C. Dong, and P. K. Liaw, Coherent precipitation and strengthening in compositionally complex alloys: A review, *Entropy* **20**, 878 (2018).
 - [7] J. Behler and M. Parrinello, Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces, *Phys. Rev. Lett.* **98**, 146401 (2007).
 - [8] D. Rogers and M. Hahn, Extended-connectivity fingerprints, *J. Chem. Inf. Model.* **50**, 742 (2010).
 - [9] A. P. Bartók and G. Csányi, Gaussian approximation potentials: A brief tutorial introduction, *Int. J. Quantum Chem.* **115**, 1051 (2015).
 - [10] F. A. Faber, L. Hutchison, B. Huang, J. Gilmer, S. S. Schoenholz, G. E. Dahl, O. Vinyals, S. Kearnes, P. F. Riley, and O. Anatole Von Lilienfeld, Prediction errors of molecular machine learning models lower than hybrid dft error, *J. Chem. Theo. Com.* **13**, 5255 (2017).
 - [11] D. Dragoni, T. D. Daff, G. Csányi, and N. Marzari, Achieving DFT accuracy with a machine-learning interatomic potential: Thermomechanics and defects in bcc ferromagnetic iron, *Phys. Rev. Mater.* **2**, 13808 (2018).
 - [12] W. J. Szlachta, A. P. Bartók, and G. Csányi, Accuracy and transferability of Gaussian approximation potential models for tungsten, *Phys. Rev. B* **90**, 104108 (2014).
 - [13] L. Zhang, D.-Y. Lin, H. Wang, R. Car, and E. Weinan, Active learning of uniformly accurate interatomic potentials for materials simulation, *Phys. Rev. Mater.* **3**, 23804 (2019).
 - [14] N. Bernstein, G. Csányi, and V. L. Deringer, De novo exploration and self-guided learning of potential-energy surfaces, *npj Comput. Mater.* **5**, (2019).
 - [15] R. Jinnouchi, F. Karsai, and G. Kresse, On-the-fly machine learning force field generation: Application to melting points, *Phys. Rev. B* **100**, 014105 (2019).
 - [16] L. Ward, R. Liu, A. Krishna, V. I. Hegde, A. Agrawal, A. Choudhary, and C. Wolverton, Including crystal structure attributes in machine learning models of formation energies via Voronoi tessellations, *Phys. Rev. B* **96**, 024104 (2017).
 - [17] T. Xie and J. C. Grossman, Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties, *Phys. Rev. Lett.* **120**, 145301 (2018).
 - [18] A. Jain and T. Bligaard, Atomic-position independent descriptor for machine learning of material properties, *Phys. Rev. B* **98**, 214112 (2018).
 - [19] F. Maresca, D. Dragoni, G. Csányi, N. Marzari, and W. A. Curtin, Screw dislocation structure and mobility in body centered cubic Fe predicted by a Gaussian Approximation Potential, *npj Comput. Mater.* **4**, 69 (2018).
 - [20] G. Imbalzano, A. Anelli, D. Giofré, S. Klees, J. Behler, and M. Ceriotti, Automatic selection of atomic fingerprints and

- reference configurations for machine-learning potentials, *J. Chem. Phys.* **148**, 241730 (2018).
- [21] G. D. Preston, Structure of age-hardened aluminium-copper alloys, *Nature* **142**, 570 (1938).
- [22] J.-F. Nie, Physical metallurgy of light alloys, *Physical Metallurgy (Fifth Edition)* (Elsevier, Oxford, 2014) pp. 2009–2156.
- [23] D. Raabe, C. C. Tasan, and E. A. Olivetti, Strategies for improving the sustainability of structural metals, *Nature* **575**, 64 (2019).
- [24] C. Wolverton and V. Ozoliņš, Entropically Favored Ordering: The Metallurgy of Al₂Cu Revisited, *Phys. Rev. Lett.* **86**, 5518 (2001).
- [25] F. Apostol and Y. Mishin, Interatomic potential for the Al-Cu system, *Phys. Rev. B* **83**, 054116 (2011).
- [26] C. V. Singh and D. H. Warner, An atomistic-based hierarchical multiscale examination of age hardening in an Al-Cu alloy, *Metall. Mater. Trans. A* **44**, 2625 (2013).
- [27] G. Esteban-Manzanares, E. Martínez, J. Segurado, L. Capolungo, and J. LLorca, An atomistic investigation of the interaction of dislocations with Guinier-Preston zones in Al-Cu alloys, *Acta Mater.* **162**, 189 (2019).
- [28] B. Wu, Z. Bai, A. Misra, and Y. Fan, Atomistic mechanism and probability determination of the cutting of Guinier-Preston zones by edge dislocations in dilute Al-Cu alloys, *Phys. Rev. Mater.* **4**, 020601 (2020).
- [29] S. P. Huber, S. Zoupanos, M. Uhrin, L. Talirz, L. Kahle, R. Häuselmann, D. Gresch, T. Müller, A. V. Yakutovich, C. W. Andersen, F. F. Ramirez, C. S. Adorf, F. Gargiulo, S. Kumbhar, E. Passaro, C. Johnston, A. Merkys, A. Cepellotti, N. Mounet, N. Marzari, B. Kozinsky, and G. Pizzi, AiiDA 1.0, a scalable computational infrastructure for automated reproducible workflows and data provenance, *Sci Data* **7**, 300 (2020).
- [30] S. Kirklin, J. E. Saal, B. Meredig, A. Thompson, J. W. Doak, M. Aykol, S. Rühl, and C. Wolverton, The Open Quantum Materials Database (OQMD): Assessing the accuracy of DFT formation energies, *npj Comput. Mater.* **1**, 15010 (2015).
- [31] S. P. Ong, W. D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V. L. Chevrier, K. A. Persson, and G. Ceder, Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis, *Comput. Mater. Sci.* **68**, 314 (2013).
- [32] A. Togo and I. Tanaka, First principles phonon calculations in materials science, *Scr. Mater.* **108**, 1 (2015).
- [33] A. B. Suriani, S. Alfarisa, A. Mohamed, A. Kamari, N. Hashim, I. M. Isa, M. H. Mamat, M. F. Malek, and M. K. Ahmad, Amorphous Al-Cu alloy nanowires decorated with carbon spheres synthesised from waste engine oil, *J. Alloys Compd.* **642**, 111 (2015).
- [34] 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, QUANTUM ESPRESSO: A modular and open-source software project for quantum simulations of materials, *J. Phys.: Condens. Matter* **21**, 395502 (2009).
- [35] J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized Gradient Approximation Made Simple, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [36] J. D. Pack and H. J. Monkhorst, Special points for Brillouin-zone integrations, *Phys. Rev. B* **16**, 1748 (1977).
- [37] M. Methfessel and A. T. Paxton, High-precision sampling for Brillouin-zone integration in metals, *Phys. Rev. B* **40**, 3616 (1989).
- [38] G. Prandini, A. Marrazzo, I. E. Castelli, N. Mounet, and N. Marzari, Precision and efficiency in solid-state pseudopotential calculations, *npj Comput. Mater.* **4**, 72 (2018).
- [39] A. D. Corso, Pseudopotentials periodic table: From H to Pu, *Comput. Mater. Sci.* **95**, 337 (2014).
- [40] G. Pizzi, A. Cepellotti, R. Sabatini, N. Marzari, and B. Kozinsky, AiiDA: Automated interactive infrastructure and database for computational science, *Comput. Mater. Sci.* **111**, 218 (2016).
- [41] T. Kluyver, B. Ragan-kelley, F. Pérez, B. Granger, M. Bussonnier, J. Frederic, K. Kelley, J. Hamrick, J. Grout, S. Corlay, P. Ivanov, D. Avila, S. Abdalla, and C. Willing, Jupyter Notebooks—A publishing format for reproducible computational workflows, in *Positioning and Power in Academic Publishing: Players, Agents and Agendas* (IOS Press, 2016), pp. 87–90.
- [42] L. Talirz, S. Kumbhar, E. Passaro, A. V. Yakutovich, V. Granata, F. Gargiulo, M. Borelli, M. Uhrin, S. P. Huber, S. Zoupanos, C. S. Adorf, C. W. Andersen, O. Schütt, C. A. Pignedoli, D. Passerone, J. VandeVondele, T. C. Schulthess, B. Smit, G. Pizzi, and N. Marzari, Materials Cloud, a platform for open computational science, *Sci Data* **7**, 299 (2020).
- [43] D. Marchand, J. Abhinav, G. Albert, and W. A. Curtin, Supplementary Data, Materials Cloud Archive: doi: 10.24435/materialscloud:d3-3r (2020).
- [44] A. Singraber, T. Morawietz, J. Behler, and C. Dellago, Parallel multistream training of high-dimensional neural network potentials, *J. Chem. Theory Comput.* **15**, 3075 (2019).
- [45] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevMaterials.4.103601> for a list of symmetry functions.
- [46] A. Singraber, J. Behler, and C. Dellago, Library-based LAMMPS implementation of high-dimensional neural network potentials, *J. Chem. Theory Comput.* **15**, 1827 (2019).
- [47] S. Plimpton, Fast parallel algorithms for short-range molecular dynamics, *J. Comput. Phys.* **117**, 1 (1995).
- [48] F. Musil, M. J. Willatt, M. A. Langovoy, and M. Ceriotti, Fast and accurate uncertainty estimation in chemical machine learning, *J. Chem. Theory Comput.* **15**, 906 (2019).
- [49] Y. Zuo, C. Chen, X.-G. Li, Z. Deng, Y. Chen, J. Behler, G. Csányi, A. V. Shapeev, A. P. Thompson, M. A. Wood, and S. Ping Ong, A performance and cost assessment of machine learning interatomic potentials, *J. Phys. Chem. A* **124**, 731 (2020).
- [50] G. P. M. Leyson, W. A. Curtin, L. G. Hector, and C. F. Woodward, Quantitative prediction of solute strengthening in aluminium alloys, *Nat. Mater.* **9**, 750 (2010).
- [51] W. A. Curtin, D. L. Olmsted, and L. G. Hector, A predictive mechanism for dynamic strain ageing in aluminium-magnesium alloys, *Nat. Mater.* **5**, 875 (2006).

- [52] V. Vaithyanathan, C. Wolverton, and L. Q. Chen, Multiscale modeling of theta' precipitation in Al-Cu binary alloys, *Acta Mater.* **52**, 2973 (2004).
- [53] B. Yin, Z. Wu, and W. A. Curtin, Comprehensive first-principles study of stable stacking faults in hcp metals, *Acta Mater.* **123**, 223 (2017).
- [54] B. Grabowski, T. Hickel, and J. Neugebauer, *Ab initio* study of the thermodynamic properties of nonmagnetic elementary fcc metals: Exchange-correlation-related error bars and chemical trends, *Phys. Rev. B* **76**, 024309 (2007).
- [55] B. Grabowski, L. Ismer, T. Hickel, and J. Neugebauer, *Ab initio* up to the melting point: Anharmonicity and vacancies in aluminum, *Phys. Rev. B* **79**, 134106 (2009).
- [56] A. Glensk, B. Grabowski, T. Hickel, and J. Neugebauer, Breakdown of the Arrhenius Law in Describing Vacancy Formation Energies: The Importance of Local Anharmonicity Revealed by *Ab Initio* Thermodynamics, *Phys. Rev. X* **4**, 011018 (2014).
- [57] A. Glensk, B. Grabowski, T. Hickel, and J. Neugebauer, Understanding Anharmonicity in fcc Materials: From its Origin to *Ab Initio* Strategies Beyond the Quasiharmonic Approximation, *Phys. Rev. Lett.* **114**, 195901 (2015).
- [58] H. K. Hardy, J. M. Silcock, and T. J. Heal, Structural Ageing Characteristics of Binary Aluminium-Copper Alloys, *J. Inst. Met.* **82**, 239 (1954).
- [59] O. I. Gorbatov, A. Yu Stroeve, Yu N. Gornostyrev, and P. A. Korzhavyi, Effective cluster interactions and pre-precipitate morphology in binary Al-based alloys, *Acta Mater.* **179**, 70 (2019).
- [60] M. Mantina, Y. Wang, L. Q. Chen, Z. K. Liu, and C. Wolverton, First principles impurity diffusion coefficients, *Acta Mater.* **57**, 4102 (2009).
- [61] H. Miyoshi, H. Kimizuka, A. Ishii, and S. Ogata, Temperature-dependent nucleation kinetics of Guinier-Preston zones in AlCu alloys: An atomistic kinetic Monte Carlo and classical nucleation theory approach, *Acta Materialia* **179**, 262 (2019).
- [62] Yi Hu, B. A. Szajewski, D. Rodney, and W. A. Curtin, Atomistic dislocation core energies and calibration of non-singular discrete dislocation dynamics, *Model. Simul. Mater. Sci. Eng.* **28**, 015005 (2020).
- [63] A. N. Stroh, Dislocations and cracks in anisotropic elasticity, *Philos. Mag.* **3**, 625 (1958).
- [64] C. Woodward, D. R. Trinkle, L. G. Hector, and D. L. Olmsted, Prediction of Dislocation Cores in Aluminum from Density Functional Theory, *Phys. Rev. Lett.* **100**, 045507 (2008).
- [65] F. Ercolessi and J. B. Adams, Interatomic potentials from first-principles calculations: The force-matching method, *Eurphys. Lett.* **26**, 583 (1994).
- [66] Y. Mishin, D. Farkas, M. J. Mehl, and D. A. Papaconstantopoulos, Interatomic potentials for monoatomic metals from experimental data and *ab initio* calculations, *Phys. Rev. B* **59**, 3393 (1999).
- [67] P. Andric and W. A. Curtin, Atomistic modeling of fracture, *Model. Simul. Mater. Sci. Eng.* **27**, 013001 (2019).
- [68] A. K. Nair, D. H. Warner, and R. G. Hennig, Coupled quantum-continuum analysis of crack tip processes in aluminum, *J. Mech. Phys. Solids* **59**, 2476 (2011).